New Space-Time Code Constructions for Two-User Multiple Access Channels

Hsiao-Feng (Francis) Lu, Roope Vehkalahti, Camilla Hollanti*,
Jyrki Lahtonen, Yi Hong, and Emanuele Viterbo

Abstract

This paper addresses the problem of constructing multiuser multiple-input multiple-output (MU-MIMO) codes for two users. The users are assumed to be equipped with n_t transmit antennas, and there are n_r antennas available at the receiving end. A general scheme is proposed and shown to achieve the optimal diversity-multiplexing gain tradeoff (DMT). Moreover, an explicit construction for the special case of $n_t = 2$ and $n_r = 2$ is given, based on the optimization of the code shape and density. All the proposed constructions are based on cyclic division algebras and their orders and take advantage of the multi-block structure. Computer simulations show that both the proposed schemes yield codes with excellent performance improving upon the best previously known codes. Finally, it is shown that the previously proposed design criteria for DMT optimal MU-MIMO codes are sufficient but in general too strict and impossible to fulfill. Relaxed alternative design criteria are then proposed and shown to be still sufficient for achieving the multiple-access channel diversity-multiplexing tradeoff.

I. Introduction

During the past five years extensive research has been carried out on single-user (SU) multiple-input multiple-output (MIMO) space-time (ST) lattice codes based on cyclic division algebras (CDAs) [1], [2], [3], [4], [5]. At its best, this research has resulted in codes that get very close to the outage bound for practical numbers of antennas. Motivated by the promising outcome in the SU-MIMO scenario, the aim in this paper is to adapt the machinery provided by CDAs to the multiuser (MU) MIMO scenario as well, with the ultimate goal of producing diversity-multiplexing tradeoff (DMT) achieving codes in mind. We will concentrate on the multiple-access channel (MAC), i.e., on the uplink transmission from

This work was supported in part by the Emil Aaltonen Foundation, Finland, through a grant for R. Vehkalahti, and by the Finnish Academy of Science and Letters and the Academy of Finland (grant #210280) through grants for C. Hollanti.

H.-F. Lu is with Department of Communications Engineering, National Chiao Tung University, Hsinchu, Taiwan (e-mail: francis@cc.nctu.edu.tw).

R. Vehkalahti, C. Hollanti (*contact person), and J. Lahtonen are with Department of Mathematics, FI-20014 University of Turku, Finland (e-mails: {roiive,cajoho,lahtonen}@utu.fi).

Y. Hong is with the Institute of Advanced Telecommunications, Swansea University, Swansea, UK. (e-mail:Y.Hong@swansea.ac.uk).

E. Viterbo is with DEIS - Universitá Della Calabria, Via P. Bucci, 42/C, 87036 Rende (CS), Italy. (e-mail:viterbo@deis.unical.it).

multiple users to a single access point (AP). Both the transmitters (=users) and the receiver (=AP) may be occupied with multiple antennas.

In general, multiuser MIMO coding is a very challenging topic. When the 3GPP (=third generation partnership project) asked the participating companies (cell phone manufacturers, chipset manufacturers, operators etc.) to list research topics that they find essential for the next release, MU-MIMO was mentioned in nearly all the lists. The area is made very challenging by the diversity of potential applications all requiring slightly different treatment and design goals.

The idea of extending the single-user ST codes to the multiuser case and the design criteria for such MU-MIMO codes were given in [6]. An explicit (2×2) two-user MIMO construction exploiting independent Alamouti blocks was also introduced in [6]. By swapping columns for one user they managed to achieve a minimum rank of three. In [7], Tse *et al.* extended the DMT results from [8] to the MAC. The codes in [6] do not achieve the optimal MAC DMT. Nam *et al.* [9] proposed the first explicit DMT achieving transmission scheme based on a class of structured multiple access lattice ST codes. However, their scheme was not constructive and no explicit examples were provided. Some explicit, algebraic code constructions for the MAC with $n_t > 1$ were introduced in [10] and [11]. The authors of [11] state that their construction is DMT optimal, but do not provide an explicit proof. In [10] a somewhat different approach was taken as compared to [6]: the authors propose a design criteria based on a truncated union-bound approximation. With the aid of these criteria they manage to outperform in error performance the other known two-user codes for the (2×2) MAC [6], [12]. Another group of multiuser ST codes was proposed in [12], but these codes suffer from high peak-to-average power ratio (PAPR) as the codeword matrices contain zero entries. In [13], the authors propose design criteria for designing MAC-DMT optimal codes, and further propose a code construction that is claimed to fulfill their criteria. The criteria proposed in [13] are indeed *sufficient* for achieving the optimal DMT, but it turns out that it is *not necessary* to fulfill these criteria in order to do so. It will be shown that more relaxed design criteria will still provide us with MAC-DMT optimal codes. Especially, we will prove that it is not possible to design DMT optimal multiuser codes having the full NVD property when we have two users using one antenna. The general proof for an arbitrary number of users and antennas is presented in [14].

Our main goals in this paper are to

1) construct explicit, sphere-decodable codes for the (2×2) situation where both of the two users are equipped with two transmitting antennas, and two antennas are available at the receiving end. We will compare our codes with the best known codes for this situation [10].
2) design a general, DMT-achieving, sphere-decodable $(n_t \times n_r)$ MU-MIMO scheme for two users, that would yield good performance also at the low SNR end. We will compare our explicit (2×4) codes with the best known codes for this situation [11].

For the use of matrix representations of cyclic division algebras and their orders as space-time codes, we refer the reader to [15], [1], [5].

Remark I.1. Our aim is to preserve maximum-likelihood (ML) performance, hence the requirement of sphere decodability. Other than sphere decoding considered here, there are still other ways to decode. For example, the MMSE-DFE combined with lattice decoding was discussed by Belfiore *et al.* in their paper [11]. Suboptimal decoders were indeed proposed for any number of receivers in [16]. Albeit in some cases they can get very close, they will always lose the ML performance at least to some extent, especially when they are necessarily needed (e.g. a sphere decoder would not work due to a lattice dimension too high). This is due to the fact that when the receiver is trying to decode a lattice that has higher dimension than the receiver's vector space, it means that there exist lattice points that are arbitrarily close to each other and hence impossible to tell apart from each other. If, in such scenario, we pursue on suboptimal tricks, some structure is then bound to be lost and the performance can never be as good as ML decoding promises, though close we get.

However, when the number of antennas increases, suboptimal decoders are the only reasonable possibility even for our (in principle) sphere decodable codes as sphere decoding will get too complex when the dimension of the lattice grows big enough.

The paper is organized as follows. In Section II we provide the reader with algebraic preliminaries, concentrating only on the facts that will be needed in this paper. Section III is devoted to designing a 2×2 two-user code, whereas Section IV gives us a general DMT optimal $n_t \times n_r$ construction for two users. In Appendix I we prove the claimed non-existence result of full-NVD multiuser codes in the case of two users equipped with one antenna.

II. Algebraic preliminaries

In this section we introduce some concepts and results from the theory of central simple algebras for later use. For the proofs of these results and for a proper introduction we refer the reader to [17].

In the rest of the paper we assume that all the fields are finite extensions of the field of rational numbers $\mathbf{Q}$.

Definition II.1. Let K be an algebraic number field and assume that E/K is a cyclic Galois extension of degree n with Galois group $\mathrm{Gal}(E/K)=\langle\sigma\rangle$. We can now define an associative K-algebra

$$\mathfrak{A}=(E/K,\sigma,\gamma)=E\oplus uE\oplus u^2E\oplus\cdots\oplus u^{n-1}E,$$

where $u\in\mathfrak{A}$ is an auxiliary generating element subject to the relations $xu=u\sigma(x)$ for all $x\in E$ and $u^n=\gamma\in K^*$. We call this type of algebra a *cyclic algebra* and the field K the *center* of the algebra. The center is the set of elements of $\mathfrak{A}$ that commute with all the elements of $\mathfrak{A}$. Throughout the paper, K denotes the center, and F denotes its subfield $F\subseteq K$. The inclusion may also be trivial, i.e., we allow $K=F$.

Definition II.2. A cyclic algebra is a division algebra if and only if all the non-zero elements of the algebra are invertible.

Proposition II.1 (Norm condition)**.** *The cyclic algebra $\mathfrak{A}=(E/K,\sigma,\gamma)$ of degree n is a division algebra if and only if the smallest factor $t\in\mathbf{Z}_+$ of n such that γ^t is the norm of some element of E^* is n.*

Due to the above proposition, the element γ is often referred to as the *non-norm element*.

Definition II.3. Let $\mathfrak{D}$ be a K-central division algebra. We then call $\sqrt{[\mathfrak{D}:K]}$ the index of the algebra.

Definition II.4. Suppose that E is a cyclic extension of an algebraic number field K. Let $\mathfrak{D}=(E/K,\sigma,\gamma)$ be a cyclic division algebra and let $\gamma \in K^*$ to be an algebraic integer. We immediately see that the $\mathscr{O}_K$-module

$$\Lambda = \mathscr{O}_E \oplus u\mathscr{O}_E \oplus \cdots \oplus u^{n-1}\mathscr{O}_E,$$

where $\mathscr{O}_E$ is the ring of integers of E, is a subring in the cyclic algebra $(E/K,\sigma,\gamma)$. We refer to this ring as the *natural order*. Note also that if γ is not an algebraic integer, then Λ fails to be closed under multiplication.

Let K/F be a finite extension (could be also the trivial extension) of algebraic number fields and $\mathfrak{D}$ a K-central division algebra of degree n.

Definition II.5. An $\mathscr{O}_F$-order Λ in $\mathfrak{D}$ is a subring of $\mathfrak{D}$, having the same identity element as $\mathfrak{D}$, and such that Λ is a finitely generated module over $\mathscr{O}_F$ and generates $\mathfrak{D}$ as a linear space over F.

Proposition II.2. *Every $\mathscr{O}_K$-order $\Lambda \subseteq \mathfrak{D}$ is also an $\mathscr{O}_F$-order.*

Definition II.6. An $\mathscr{O}_F$-order Λ is called *maximal*, if it is not properly contained in any other $\mathscr{O}_F$-order.

Proposition II.3. *Any K-central division algebra $\mathfrak{D}$ has a maximal $\mathscr{O}_F$-order and any order inside $\mathfrak{D}$ is contained in at least one maximal order.*

Example II.1. Suppose that E/K is a cyclic extension of algebraic number fields. Let $\mathfrak{D}=(E/K,\sigma,\gamma)$ be a cyclic algebra.

We can consider $\mathfrak{D}$ as a right vector space over E, and every element $a = x_0 + ux_1 + \cdots + u^{n-1}x_{n-1} \in \mathfrak{D}$ has the following representation as a matrix

$$A = \begin{pmatrix} x_0 & \gamma\sigma(x_{n-1}) & \gamma\sigma^2(x_{n-2}) & \cdots & \gamma\sigma^{n-1}(x_1) \\ x_1 & \sigma(x_0) & \gamma\sigma^2(x_{n-1}) & & \gamma\sigma^{n-1}(x_2) \\ x_2 & \sigma(x_1) & \sigma^2(x_0) & & \gamma\sigma^{n-1}(x_3) \\ \vdots & & & & \vdots \\ x_{n-1} & \sigma(x_{n-2}) & \sigma^2(x_{n-3}) & \cdots & \sigma^{n-1}(x_0) \end{pmatrix}.$$

We call this representation the *left regular representation* and denote $A=\psi(a)$.

Definition II.7. The determinant (resp. trace) of the matrix A above is called the *reduced norm* (resp. *reduced trace*) of the element $a \in \mathfrak{D}$ and is denoted by $nr_{\mathfrak{D}/K}(a)$ (resp. $tr_{\mathfrak{D}/K}(a)$).

Proposition II.4. *Let $\mathfrak{D}$ be a K-central division algebra and a an element of $\mathfrak{D}$. Then $nr(a)$* and *$tr(a) \in K$.*

Proposition II.5. *The norm and trace maps do not depend on the maximal representation, i.e., the left regular representation is not the only representation we can use. However, we stick to ψ for simplicity.*

Definition II.8. We then define the reduced trace and norm of a to F by

$$tr_{\mathfrak{D}/F}(a) = tr_{K/F}(tr_{\mathfrak{D}/K}(a)) \text{ and } nr_{\mathfrak{D}/F}(a) = nr_{K/F}(nr_{\mathfrak{D}/K}(a)),$$

where $nr_{K/F}$ and $tr_{K/F}$ are the usual relative norm and trace maps of a number field extension (sometimes also denoted by $\mathrm{N}_{K/F}$ and $\mathrm{T}_{K/F}$).

Proposition II.6. *Let Λ be an $\mathcal{O}_F$-order in a K-central division algebra $\mathfrak{D}$. Then for any element $a \in \Lambda$ its reduced norm $nr_{\mathfrak{D}/F}(a)$ and reduced trace $tr_{\mathfrak{D}/F}(a)$ are elements of the ring of integers $\mathcal{O}_F$ of the field F. If a is non-zero, then so is $nr_{\mathfrak{D}/F}(a)$.*

Now we are ready to define one of the main algebraic objects needed in this paper.

Definition II.9. Let $\mathfrak{D}$ be a K-central division algebra and $m = \dim_F \mathfrak{D}$. The *$\mathcal{O}_F$-discriminant* of the $\mathcal{O}_F$-order Λ is the ideal $d(\Lambda/\mathcal{O}_F)$ in $\mathcal{O}_F$ generated by the set

$$\{\det(tr_{\mathfrak{D}/F}(x_i x_j))_{i,j=1}^{m} \mid (x_1, ..., x_m) \in \Lambda^m\}.$$

Here $\dim_F \mathfrak{D}$ simply refers to the dimension of $\mathfrak{D}$ as an F-linear vector space. If Λ is a free $\mathcal{O}_F$-module, then

$$d(\Lambda/\mathcal{O}_F) = \det(tr(x_i x_j))_{i,j=1}^{m},$$

where $\{x_1, \ldots, x_m\}$ is any $\mathcal{O}_F$-basis of Λ.

Proposition II.7. *All the maximal orders of a K-central division algebra share the same discriminant.*

Now we can define the following.

Definition II.10. Let $\mathfrak{D}$ be a K-central division algebra and let Λ be some maximal order in $\mathfrak{D}$. Then we refer to $d(\Lambda/\mathcal{O}_K) = d_{\mathfrak{D}}$ as the *discriminant of the algebra* $\mathfrak{D}$.

The following lemma connects the discriminants $d(\Lambda/\mathcal{O}_K)$ and $d(\Lambda/\mathcal{O}_F)$.

Lemma II.8. *Let $\mathfrak{D}$ be a K-central division algebra of index n and let Λ be an $\mathcal{O}_K$-order. If Λ is an $\mathcal{O}_F$-order in $\mathfrak{D}$, then*

$$d(\Lambda/\mathcal{O}_F) = nr_{K/F}(d(\Lambda/\mathcal{O}_K))d(\mathcal{O}_K/\mathcal{O}_F)^{n^2}.$$

III. A Sphere decodable MU-MIMO code for two users and two receive antennas

In this section we concentrate on designing a multiuser code for two users, both equipped with two transmit antennas, and for a receiver that has two antennas. This leads us to a situation where the single user must use a code that is sphere decodable with one receive antenna. Such MU-MIMO codes have been considered by Gärtner and Bölcskei [6] and by Hong and Viterbo in [10]. Our coding scheme is directly comparable to their codes.

In what follows, we first concentrate on the optimization of the single user code and then, in the very end of this section, we put our single-user codes into use in the multiuser scenario. The careful construction of the single-user code as a building block of the multiuser code is crucial, as it will then guarantee good performance also when only one user is present.

A. Coding theoretic preliminaries of abstract multi-block codes

In this section we consider abstract multi-block codes that are matrix lattices in the space $M_{n\times nk}(\mathbf{C})$. Particularly we are going to define the *normalized minimum determinant* and *normalized coding gain* of such lattices and study the relation between these concepts.

We can flatten the matrices A of $M_{n\times nk}(\mathbf{C})$ to real vectors $\alpha(A) \in \mathbf{R}^{2kn^2}$ by first forming a vector of length kn^2 out of the entries (e.g. row by row) and then replacing a complex number z with the pair of its real and imaginary parts $\Re z$ and $\Im z$. This mapping α is clearly $\mathbf{R}$-linear and maps t-dimensional $M_{n\times nk}(\mathbf{C})$ lattices to t-dimensional $\mathbf{R}^{2kn^2}$ lattices. We also have the equality $||A||_F = ||\alpha(A)||_E$, i.e., the Frobenius norm of the matrix A coincides with the euclidean norm of the corresponding vector $\alpha(A)$. Therefore, α is also an isometry.

Definition III.1. We say that a lattice $\mathbf{L}$ in $M_{n\times nk}(\mathbf{C})$ is *orthogonal* or *rectangular* if the corresponding real lattice $\alpha(\mathbf{L})$ has a basis that is orthogonal with respect to the normal inner product of the space $\mathbf{R}^{2kn^2}$.

We denote the measure (or hypervolume) of the fundamental parallelotope of the lattice $\alpha(\mathbf{L})$ by $m(\mathbf{L})$ and we call it the *volume of the fundamental parallelotope of the lattice* $\mathbf{L}$. If $\{x_1,\dots,x_t\}$ is a basis of $\mathbf{L}$, we can form a matrix M by using the vectors $\alpha(x_i)$ as column blocks. Then the *Gram matrix* of the lattice $\mathbf{L}$ is

$$G(\mathbf{L}) = MM^T = \left(\Re tr(x_i x_j^\dagger)\right)_{1\leq i,j\leq t},$$

where $X^\dagger$ indicates the complex conjugate transpose of X. The Gram matrix then has a positive determinant equal to $m(\mathbf{L})^2$.

Any lattice $\mathbf{L} \subseteq M_{n\times nk}(\mathbf{C})$ can be scaled (i.e. multiplied by a real constant s) to satisfy $m(sL) = 1$.

If A is an element in the space $M_{n\times nk}(\mathbf{C})$ it can be written as $(A_1,\dots,A_k)$ where all the matrices A_i are elements in $M_{n\times n}$. We can then define the *product determinant*

$$\mathrm{pdet}(A) = \prod_{i=1}^{k} \det(A_i)$$

of the matrix A.

Definition III.2. The minimum determinant $\det_{min}(\mathbf{L})$ of a multi-block code $\mathbf{L} \subseteq M_{n\times nk}(\mathbf{C})$ is defined to be the infimum of the absolute values $\mathrm{pdet}(A)$ of all the non-zero elements of the lattice $\mathbf{L}$.

The *normalized minimum determinant* $\delta(\mathbf{L})$ of a lattice $\mathbf{L}$ is obtained by multiplying the lattice with a real constant such that the resulting lattice $\mathbf{L}'$ has fundamental parallelotope of volume 1 and then setting

$$\delta(\mathbf{L}) = \det_{min}\left(\mathbf{L}'\right).$$

Definition III.3. The coding gain $CG(\mathbf{L})$ of the lattice $\mathbf{L} \subseteq M_{n\times nk}(\mathbf{C})$, $k \geq n$, is defined to be the infimum of the absolute values of the determinants of matrices $AA^\dagger$ of all non-zero matrices A in the lattice.

The *normalized coding gain* $NCG(\mathbf{L})$ of a lattice $\mathbf{L} \subseteq M_{n\times nk}(\mathbf{C})$ is obtained by multiplying the lattice by a real constant such that the resulting lattice $\mathbf{L}'$ has a fundamental parallelotope of volume 1 and then set

$$NCG(\mathbf{L}) = CG(\mathbf{L}').$$

Lemma III.1. *Let us suppose that $A_1,\dots,A_k$ are complex $n\times n$ matrices. We consider the $n\times nk$ matrix $(A_1,A_2,\dots,A_k)=A$. We then have* $\det(AA^\dagger) \geq k^n \cdot (\prod_{i=1}^k |\det(A_i)|)^{2/k}$.

Proof: First the Minkwoski determinant inequality states that $(\det(AA^\dagger))^{(1/n)} \geq \sum_{i=1}^k |\det(A_i)|^{2/n}$. The AM-GM inequality on the arithmetic and geometric means then transforms this result into

$$\det(AA^\dagger)^{1/n} \geq \sum_{i=1}^{k} |\det(A_i)|^{2/n} \geq k \cdot (\prod_{i=1}^{k} |\det(A_i)|^{2/n})^{1/k}.$$

□

In the following corollary we use the notation of the previous lemma.

Corollary III.2. *Let us suppose that $\mathbf{L}$ is a multi-block code in $M_{n\times nk}(\mathbf{C})$. Then*

$$CG(\mathbf{L}) \geq k^n(\det_{min}(\mathbf{L}))^{2/k} \quad \text{and} \quad NCG(\mathbf{L}) \geq k^n(\delta(\mathbf{L}))^{2/k}.$$

Particularly the following will be of great interest for us.

Corollary III.3. *Let us suppose that $\mathbf{L}$ is a lattice in $M_{2\times 4}(\mathbf{C})$. Then $NCG(\mathbf{L}) \geq 2^2\delta(\mathbf{L})$.*

Remark III.1. The concept of the normalized minimum determinant of a multi-block code is related to the performance of the code when each $n\times n$ block faces independent fading. On the other hand, the normalized coding gain is a relevant code design criterion when the channel stays stable during the transmission of the whole $n\times nk$ block. It is not a great surprise that these two concepts are so closely related.

B. Constructing the single user code

In this section we study the achievable normalized minimum determinant of 8-dimensional multi-block codes in the space $M_{2\times 4}(\mathbf{C})$. Notice that as we want to receive with only two antennas (equipped with sphere decoders), we *cannot use full lattices* that would have dimension 16. In order to get well behaving 8-dimensional lattices we use real quadratic field as a center in the multi-block construction. We remark that while we came up with the idea independently it was discovered already in [18].

We begin by considering maximal order codes from division algebras. By discriminant analysis we are able to find the optimal algebras. In Section III-D we concentrate on rectangular codes and derive a bound for normalized minimum determinant of such codes and give an example code achieving this bound. The minimum determinant analysis we are using is similar to that used in [19].

We will take advantage of multi-block constructions from division algebras. In Section IV to follow the same trick will be used. The exception is that now the base field F is $\mathbf{Q}$ and the center K is some quadratic field, whereas in Section IV we need full lattices; hence $F = \mathbf{Q}(i)$ and the center K is some suitable extension of F.

Let us consider the field $E = KL$ that is a compositum of two quadratic fields K and L. We suppose that $K \cap L = \mathbf{Q}$ and that $\mathrm{Gal}(K/\mathbf{Q}) =< \tau >$ and $\mathrm{Gal}(L/\mathbf{Q}) =< \sigma >$. We can then write that $\mathrm{Gal}(E/\mathbf{Q}) =< \sigma > \otimes < \tau >$.

Let us now consider the cyclic division algebra $\mathfrak{D} = (E/K, \sigma, \gamma)$. As usually, we have the left regular representation ψ of the algebra $\mathfrak{D}$ so that an element a maps to a 2×2 matrix $\psi(a) \in M_2(E)$, and the multi-block representation ϕ;

$$\phi(a) \mapsto (\psi(a), \tau(\psi(a))). \tag{1}$$

Let us suppose that Λ is a $\mathbf{Z}$-order in $\mathfrak{D}$. We call the $\phi(\Lambda)$ an *order code*. In the rest of this section, we suppose that the division algebras under consideration are of the previous type.

Lemma III.4. *Let a be an element of* $\mathfrak{D}$. *Then*

$$\det(\psi(a))\det((\tau(\psi(a))) = nr_{\mathfrak{D}/\mathbf{Q}}(a) \quad \text{and} \quad \mathrm{Tr}(\psi(a) + \tau(\psi(a)) = tr_{\mathfrak{D}/\mathbf{Q}}(a),$$

where Tr is the usual matrix trace.

Proof: These results follow directly from Definition II.8. □

Proposition III.5. *Let us suppose that* Λ *is a* $\mathbf{Z}$*-order of a division algebra* $\mathfrak{D}$ *and that* ϕ *is a multi-block representation. The order code* $\phi(\Lambda)$ *is an* 8*-dimensional lattice in the space* $M_{2\times 4}(\mathbf{C})$ *and*

$$\det_{min}(\phi(\Lambda)) = 1.$$

Proof: The claim about the dimension of the lattice is easily seen. The second claim follows directly from Proposition II.6. □

Remark III.2. For every non-zero element $(\psi(a), \tau(\psi(a)))$ of an order code the rows are linearly independent over $\mathbf{C}$. This follows as $\det(\psi(a)) \neq 0$ and therefore the first two columns are linearly independent and generally in a matrix the number of linearly independent rows and columns is equal.

Corollary III.6. *With the previous notation we have* $\delta(\phi(\Lambda)) = \frac{1}{m(\phi(\Lambda))^{1/2}}$.

The previous proposition reveals that the minimum determinant of an order code depends only on the volume of the fundamental parallelotope. The following lemma connects the volume of the fundamental parallelotope and the discriminant of the algebra. Here we identify the ideal discriminant and the element generating it. This allows us to discuss the absolute value of the $\mathbf{Z}$-discriminant.

In the following we identify the order of the algebra and its image in $M_{2\times 4}(\mathbf{C})$. If the regular representation ψ of the algebra fulfills the following conditions, then the discriminant and the fundamental parallelotope of an order are tightly connected.

In the case of a real center we must assume that the regular representation ψ gives us matrices of the following Alamouti-like type

$$\begin{pmatrix} a & -b^* \\ b & a^* \end{pmatrix}, \tag{2}$$

where * is the complex conjugation. In the case of a complex center we must assume that the automorphism τ is the complex conjugation. It is an easy task to check that, with these assumptions,

$$\mathrm{Tr}(\psi(a)\psi(b)^\dagger + \tau(\psi(a))\tau(\psi(b))^\dagger) \in \mathbf{R},$$

when $a, b \in \mathfrak{D}$.

If the representation ψ fulfills the conditions stated above, then we have the following.

Lemma III.7. *Let us suppose that $\mathfrak{D}$ is a division algebra and Λ is an order in $\mathfrak{D}$. Then*

$$m(\Lambda) = \sqrt{|d(\Lambda/\mathbf{Z})|} \ \ \textit{and} \ \ \delta(\Lambda) = \frac{1}{|d(\Lambda/\mathbf{Z})|^{1/4}}.$$

Proof: Let us suppose that Λ has a $\mathbf{Z}$-basis $\mathscr{B} = \{(A_1, \tau(A_1)), \ldots, (A_8, \tau(A_8))\}$, where $A_i = \psi(a_i)$, $a_i \in \Lambda$. We can now flatten the matrix $(A_i, \tau(A_i))$ into an 8-tuple $L(A_i, \tau(A_i))$ by first forming a vector of length 4 out of the entries of A_i (e.g. row by row) and then concatenating this with the 4-tuple similarly made out of the entries of the matrix $\tau(A_i)$. We can now easily see the identities

$$L(A_i, \tau(A_i))L(A_j, \tau(A_j))^T = \mathrm{Tr}(A_i A_j^T + \tau(A_i)\tau(A_j)^T) \tag{3}$$

and

$$L(A_i, \tau(A_i))L(A_j^T, \tau(A_j)^T)^T = \mathrm{Tr}(A_i A_j + \tau(A_i)\tau(A_j)). \tag{4}$$

The Gram matrix of the lattice Λ is

$$G = (\Re(\mathrm{Tr}(A_i A_j^\dagger + \tau(A_i)\tau(A_j)^\dagger)))_{i,j=1}^8.$$

Due to the limitations we set above on the form of the matrices A_i, $\mathrm{Tr}(A_i A_j^\dagger + \tau(A_i)\tau(A_j)^\dagger)$ is already real and we can ignore taking the real part from the traces. According to Equation (3) we can write

$$G = (L(A_i, \tau(A_i))L(A_j^*, \tau(A_j)^*)^T))_{i,j=1}^8 = L(\mathscr{B})L(\mathscr{B})^\dagger,$$

where the rows of the 8×8 matrix $L(\mathscr{B})$ consist of vectors $L(A_i, \tau(A_i))$. A simple permutation of the columns and elementary properties of determinants give us that

$$|\det(L(\mathscr{B}))\det(L(\mathscr{B})^\dagger)| = |\det(L(\mathscr{B}))\det(L(\mathscr{B})^T)| = |\det(L(\mathscr{B}))\det(L(\mathscr{B}')^T)|,$$

where $L(\mathscr{B}')$ is a matrix with the rows $L((A_i)^T, \tau(A_i)^T)$. According to Equation (4) and Lemma III.4

$$L(\mathscr{B})L(\mathscr{B}')^T = (\mathrm{Tr}(A_i A_j + \tau(A_i)\tau(A_j))_{i,j=1}^8 = d(\Lambda/\mathbf{Z}).$$

□

Proposition III.8. *Of all the orders in a K-central division algebra, the maximal orders have the smallest $\mathbf{Z}$-discriminant.*

Lemma III.9. *Let us suppose that Λ is an order in a division algebra $\mathfrak{D}$. Then*

$$NCG(\Lambda) = 2^2(\delta(\Lambda))^2.$$

Proof: Let us consider the lattice Λ without the normalization. We then have $CG(\Lambda) \geq 2^2(\det_{min}(\Lambda))^2 = 2^2$. On the other hand, $\det(\phi(1_{\mathfrak{D}})\phi(1_{\mathfrak{D}})^H) = 2^2$ and therefore $CG(\Lambda) = 2^2 = 2^2(\det_{min}(\Lambda))^2$. The scaling does not destroy this equality. □

C. Minimizing the discriminant

As previously stated, if we consider orders inside a fixed algebra, the smallest discriminant belongs to the maximal orders of the algebra and all the maximal orders share the same discriminant. Among those algebras having a regular representation fulfilling the conditions stated before Lemma III.7, minimizing the discriminant of the algebra is now seen to be equivalent to maximizing the coding gain of a code from a maximal order.

In the following we forget the restrictions on the form of the regular representation and simply concentrate on finding the division algebras with the smallest possible discriminants. Only after this we shall discuss whether the algebras have such regular representations that Lemma III.7 would be at their disposal. Still the solution to the problem of choosing an optimal division algebra is not an obvious one. The first step is the following. In our special case, Lemma II.8 transforms into

$$|d(\Lambda/\mathbf{Z})| = |nr_{K/\mathbf{Q}}(d(\Lambda/\mathcal{O}_K))| d(\mathcal{O}_K/\mathbf{Z})^4.$$

Here we see that for a fixed center K the second term $d(\mathcal{O}_K/\mathbf{Z})^4$ is independent on the chosen algebra and we can concentrate on the term $|nr_{K/\mathbf{Q}}(d(\Lambda/\mathcal{O}_K))|$. This leads us to discuss the size of the ideals of $\mathcal{O}_K$. By this we mean that ideals are ordered by the absolute values of their norms to $\mathbf{Q}$, so e.g. in the case $\mathcal{O}_K = \mathbf{Z}[i]$ we say that the prime ideal generated by $2+i$ is smaller than the prime ideal generated by 3 as they have norms 5 and 9, respectively.

We have divided this section into two parts depending on the type of the center. Propositions III.10 and III.12 that consider discriminants of division algebras are straightforward corollaries of well known results and the proofs can be found for example from [17]. The minimization problems that will have rather simple solutions here become more complicated in the case where the index of the algebra is greater than two. This question is of major importance when we consider general MIMO codes. We refer the interested reader to [5].

1) A complex quadratic center: In this section we consider the situation where the center K is a complex quadratic field of degree 2.

Proposition III.10. *Let us suppose that $\mathfrak{D}$ is a K-central division algebra of index 2 containing an $\mathcal{O}_K$-order $\Lambda \subseteq \mathfrak{D}$. Then*

$$d(\Lambda/\mathcal{O}_K) = (P_1 \cdots P_{2n})^2,$$

where all the P_i are distinct prime ideals of the center K and $n \geq 1$.

On the other hand, if we have an even numbered set of prime ideals $P_1,\dots,P_{2k}$, then there exists a unique K-central division algebra $\mathfrak{D}'$ of index 2 having an $\mathscr{O}_K$-order Λ with the discriminant

$$d(\Lambda/\mathscr{O}_K) = (P_1\cdots P_{2k})^2.$$

Corollary III.11. *Suppose that P_1 and P_2 are a pair of smallest primes in the complex quadratic field K. Then the smallest $\mathbf{Z}$-discriminant of all the index 2 K-central division algebras is*

$$|nr_{K/\mathbf{Q}}(P_1P_2)|^2 d(\mathscr{O}_K/\mathbf{Z})^4.$$

Example III.1. Let us consider the center $\mathbf{Q}(i)$. It is readily seen that $(2+i)$ and $(1+i)$ are a pair of the smallest primes in this field. Proposition III.10 proves that there exists a $\mathbf{Q}(i)$-central division algebra $\mathfrak{D}$ of index 2 having a maximal order Λ with the discriminant

$$d_{\mathfrak{D}} = d(\Lambda/\mathbf{Z}) = |(1+i)(2+i)|^2 4^4 = 2^9 5.$$

If this algebra also has a suitable regular representation, then Lemma III.7 infers that

$$\delta(\Lambda) = \frac{1}{(2^9 5)^{1/4}} = 0.140....$$

Example III.2. Let us next consider the center $K=\mathbf{Q}(\sqrt{-3})$. The smallest prime ideals in this center are 2 and $\sqrt{-3}$. According to Proposition III.10 there exists a $\mathbf{Q}(\sqrt{-3})$-central division algebra $\mathfrak{D}$ of index 2 having a maximal order with the discriminant

$$d_{\mathfrak{D}} = d(\Lambda/\mathbf{Z}) = |2\sqrt{3}|^2 3^4 = 972.$$

If this algebra also has a suitable regular representation, then Lemma III.7 gives us that

$$\delta(\Lambda) = \frac{1}{(972)^{1/4}} = 0.179....$$

The discriminant 972 is already the smallest possible value we can achieve with a complex quadratic center K. This can be proved by simply trying different centers. It is easily done because for a given discriminant there is only one complex quadratic field. In the discriminant formula for the maximal order of a division algebra the term $d(\mathscr{O}_K/\mathbf{Z})^4$ is always a factor and we already have $6^4 = 1296$. Therefore it is enough to check the remaining discriminants -4 and -5 that are still possible. In the previous example we saw that the center corresponding to discriminant -4 is $\mathbf{Q}(i)$ and that with this center the discriminant cannot be smaller than 972. The discriminant of the field $\mathbf{Q}(\sqrt{-5})$ is -40 and there does not exist a field with discriminant -5.

2) A real quadratic center: In this section we fix the center K to be a real quadratic field of degree 2.

Proposition III.12. *Let us suppose that $\mathfrak{D}$ is a K-central division algebra of index 2 and that Λ is a maximal $\mathbf{Z}$-order in $\mathfrak{D}$. Then*

$$d(\Lambda/\mathscr{O}_K) = (P_1\cdots P_n)^2,$$

where P_i are separate prime ideals of K and $n \geq 0$. Here we use the notation that if $n=0$ then $d(\Lambda/\mathscr{O}_K) = \mathscr{O}_K$.

On the other hand if we have a set of prime ideal $P_1,\dots,P_k$ then there exists a K-central division algebra $\mathfrak{D}'$ of index 2 having a maximal order Λ' with discriminant

$$d(\Lambda'/\mathscr{O}_K) = (P_1\cdots P_k)^2$$

with the notation that if $k=0$, $d(\Lambda'/K)=\mathscr{O}_K$.

Corollary III.13. *Let us suppose that we have a real quadratic field K. Then the smallest discriminant of all the index 2 division algebras with the center K is*

$$d(\mathscr{O}_K/\mathbf{Z})^4.$$

Example III.3. The smallest discriminant of all the real quadratic fields belongs to the field $\mathbf{Q}(\sqrt{5})=K$. The following algebra

$$\mathfrak{D}_{icos} = (\mathbf{Q}(i,\sqrt{5})/\mathbf{Q}(\sqrt{5}),\sigma,-1)$$

is called the *Icosian algebra*. It is a known fact that $|d_{\mathfrak{D}_{icos}}|=1$. This reveals that this division algebra has the smallest $\mathbf{Z}$-discriminant of all the index two division algebras with a real quadratic center. Lemma II.8 then gives us that $d(\Lambda/\mathbf{Z})=5^4$. We immediately see that the regular presentation attached to the cyclic presentation of $\mathfrak{D}_{icos}$ fulfills the expectations of Equation 2. According to Lemma III.7 we then have that $m(\Lambda)=25$, and according to Lemma III.6

$$\delta(\Lambda)=\frac{1}{5}=0.2.$$

A comparison to complex centers proves that this algebra has the smallest discriminant of all the index two algebras where the center is a quadratic field.

Remark III.3. We remark that the order code promised to exist by the previous example actually played part in the construction of the Icosian code in [20].

The previous example gave us an idea of the achievable coding gain with order theoretic methods. Yet a simple modulation scheme can easily ruin the performance of such codes. For instance, if we use a $\mathbf{Z}$-module basis together with a PAM scheme the promised minimum determinant advantage might never get realized. Therefore the next section is devoted for constructing a code with rectangular shaping.

D. A rectangular MISO code with the best achievable minimum determinant

In this section we concentrate on the question of achievable minimum determinant of rectangular multi-block codes in the space $M_{2\times 4}(\mathbf{C})$.

Proposition III.14. *Let us suppose that* $\mathbf{L}$ *is a rectangular multi-block code in the space* $M_{2\times 4}(\mathbf{C})$. *We then have that*

$$\delta(\mathbf{L})\leq\frac{1}{16}.$$

Proof: We expect w.l.o.g. that $\mathbf{L}$ has a fundamental parallelotope of volume 1. Consider an orthogonal basis of $\mathbf{L}$. Due to the orthogonal shape at least one of the basis vectors must have length less than or equal to one. Let us suppose that $(A_1, A_2) = A$ is a matrix corresponding to such vector. This means that $||A||_F \leq 1$. Let us consider the matrix $B = \mathrm{diag}(A_1, A_2)$. According to Hadamard inequality we have that

$$|\det(A_1)\det(A_2)| = |\det(B)| \leq \frac{(||B||_F)^4}{16} = \frac{(||A||_F)^4}{16} \leq \frac{1}{16}.$$

□

In the following we are going to build an orthogonal order code that reaches the bound of the previous proposition. Let us consider the following algebra

$$\mathfrak{D}_{ort} = (\mathbf{Q}(i, \sqrt{2})/\mathbf{Q}(\sqrt{2}), \sigma, -1),$$

and the natural order Λ_{ort} of this algebra. The field $L = \mathbf{Q}(i, \sqrt{2})$ can be seen as a $\mathbf{Z}[i]$-module with a basis $\{1, \zeta_8\}$. Now the natural order can be written as

$$\Lambda_{ort} = \mathbf{Z}[i] \oplus \mathbf{Z}[i]\zeta_8 \oplus u\mathbf{Z}[i] \oplus u\mathbf{Z}[i]\zeta_8.$$

The operation of the automorphism τ is defined as $\tau(\zeta_8) = -\zeta_8$, $\tau(i) = i$ and σ is just the usual complex conjugation. The multi-block representation ϕ now gives us that

$$\phi(a_1 + a_2\zeta_8 + ua_3 + u\zeta_8 a_4) =$$
$$\begin{pmatrix} (a_1 + a_2\zeta_8) & -\overline{(a_3 + a_4\zeta_8)} & a_1 - a_2\zeta_8 & -\overline{(a_3 - a_4\zeta_8)} \\ (a_3 + a_4\zeta_8) & \overline{(a_1 + a_2\zeta_8)} & a_3 - a_4\zeta_8 & \overline{a_1 - a_2\zeta_8} \end{pmatrix}.$$

By simply checking we see that

$$\{1, i, \zeta_8, \zeta_8 i, u, ui, u\zeta_8, u\zeta_8 i\}$$

forms a rectangular basis for the code. A particularly nice feature of this code is that we can apply QAM-modulation here, although the general construction method did not promise this.

We could now just calculate the fundamental parallelotope of this code and then determine the normalized minimum determinant, but we take a more general approach that sheds more light to the question of how we first came up with this code.

Lemma III.15. *[5, Lemma 2.9] Let us suppose that K is such an algebraic number field that $\mathcal{O}_K$ is a principal ideal domain. If $\mathfrak{D} = (E/K, \sigma, \gamma)$ is a K-central division algebra of index n and Λ is a natural order in $\mathfrak{D}$, then*

$$|d(\Lambda/\mathbf{Z})| = |d(E/\mathbf{Q})^n \gamma^{2n(n-1)}|.$$

We now return to our example algebra above and to the fixed natural order Λ_{ort} in it. The discriminant of the extension $\mathbf{Q}(i, \sqrt{2})/\mathbf{Q}$ has absolute value 256. Lemma III.15 now states that

$$|d(\Lambda_{ort}/\mathbf{Z})| = 256^2$$

and because the left regular representation in this case is suitable Lemma III.7 gives us that

$$\delta(\Lambda_{ort}) = \frac{1}{16}.$$

Remark III.4. The code Λ_{ort} appeared in [21] as a 4×1 MISO code. It was noted that Λ_{ort} is unitarily equivalent to their L_2 code.

E. A multiuser coding scheme

In this section we propose a simple multiuser coding scheme that is based on our previous work on MISO codes. The scheme is based on the criteria presented in [6].

As an example we apply the code of Section III-D and compare its performance to the corresponding codes in [6] and [10].

Let us assume that

$$\Gamma = \begin{pmatrix} \zeta & 0 \\ 0 & \zeta \end{pmatrix},$$

where ζ is some primitive m^{th} root of unity, m being sufficiently large so that ζ cannot possibly be a root for the determinant polynomial, meaning that our 2-user code matrix will end up having rank 4. If only one user is transmitting the situation is equal to delay four 2×2 single-user MIMO transmission.

The infinite code lattice for the first user is $\alpha(\Lambda)$ where

$$\alpha(a) = \begin{pmatrix} \Gamma\psi(a) & \tau(\psi(a)) \end{pmatrix},$$

where $a \in \Lambda$. The single user code lattice for the second user is $\beta(\Lambda)$, where

$$\beta(b) = \begin{pmatrix} \psi(b) & \Gamma\tau(\psi(b)) \end{pmatrix},$$

and $b \in \Lambda$.

If the users are independent yet synchronized the signal sent by the two users is

$$C = \begin{pmatrix} \alpha(a) \\ \beta(b) \end{pmatrix}.$$

If we suppose that neither a or b is zero, then the determinant of the matrix C is a polynomial of ζ and the term attached to its highest power is $\psi(a)\tau(\psi(b))$. By our assumption this term is non-zero. If ζ is now a suitable primitive m^{th} root of unity, we see that as long as a and b are non-zero elements, matrix C has rank 4. If only one user is transmitting, then by Remark III.2 the matrix has rank 2.

Let us now consider a sample code based on our orthogonal code of Section III-D.

The code for the first user is

$$\begin{pmatrix} \zeta_7(a_1+a_2\zeta_8) & \zeta_7(-\overline{(a_3+a_4\zeta_8)}) & a_1-a_2\zeta_8 & -\overline{(a_3-a_4\zeta_8)} \\ \zeta_7(a_3+a_4\zeta_8) & \zeta_7\overline{(a_1+a_2\zeta_8)} & a_3-a_4\zeta_8 & \overline{a_1-a_2\zeta_8} \end{pmatrix}$$

and for the second user

$$\begin{pmatrix} b_1+b_2\zeta_8 & -\overline{(b_3+b_4\zeta_8)} & \zeta_7(b_1-b_2\zeta_8) & \zeta_7(-\overline{(b_3-b_4\zeta_8)}) \\ b_3+b_4\zeta_8 & \overline{b_1+b_2\zeta_8} & \zeta_7(b_3-b_4\zeta_8) & \zeta_7\overline{(b_1-b_2\zeta_8)} \end{pmatrix},$$

where a_i and b_i are QAM-symbols.

F. Simulations

In this section we compare our code construction to two previously proposed codes [10] (HV) and [6] (GB).

In [6] the coding scheme consist of two single user codes

$$U_1 = \begin{pmatrix} x_1(1) & x_1(2) & x_1(3) & x_2(4) \\ x_1(2)^* & x_1(1) & x_2(4)^* & x_1(3)^* \end{pmatrix} \text{ and } U_2 = \begin{pmatrix} x_2(1) & x_2(3) & x_2(2) & x_2(4) \\ x_2^*(1) & x_2(4)^* & x_2(1) & x_2(3)^* \end{pmatrix},$$

where in both cases the symbols $x_i(j)$ are independently chosen from some QAM-constellation. When both users are transmitting the combined matrix has rank 3 (see [6]).

In [10] the HV code is based on the number field code used in the construction of the 4×4 Perfect code [1]. The key parts are the field extension $L/K = \mathbf{Q}(i, \zeta_{15} + \zeta_{15})/\mathbf{Q}(i)$, its cyclic Galois group $G(L/K) = < \sigma >$, and an ideal I of the ring of algebraic integers $\mathscr{O}_L$. Here the single user codes are

$$U_1 = \begin{pmatrix} a & \sigma(a) & \sigma^2(a) & \sigma^3(a) \\ i\sigma^3(a) & \sigma(a)^2 & \sigma(a) & a \end{pmatrix} \text{ and } U_2 = \begin{pmatrix} ib & i\sigma(b) & \sigma^2(b) & \sigma^3(b) \\ i\sigma^3(a) & i\sigma(a)^2 & i\sigma(a) & a \end{pmatrix},$$

where a and b are elements of the ideal I corresponding to a given QAM constellation. When both users are transmitting the combined 4×4 matrix has rank 4, and when only one user is transmitting the rank is 2 (see [10]).

In Figures 1 and 2 we compare our new code (NC) to the codes in [10] (HV) and [6] (GB) in a slow fading situation where the channel remains fixed for four channel uses. We see a considerable gain compared to the previous code constructions. When compared to the GB code the performance advantage is explained by the fact that when both users are transmitting, the combined matrix of the NC code has rank 4, whereas the GB code has rank 3 only. Both codes are taking full advantage of the delay four, but encoding of the GB code is perhaps simpler. The decoding of both the GB code and the NC code can be simply done using a sphere decoder. Both the GB code and the NC code involve an Alamouti-like structure which can be taken advantage of in the decoding process.

When comparing the HV code and the NC code we have tie on ranks, but the optimality of our single user codes (see Proposition III.14) expectedly gives us an edge in coding gain. In this case the encoding and decoding processes have similar complexity.

IV. DMT OPTIMAL CODE CONSTRUCTION FOR TWO USERS

In this section we will focus on the construction of DMT optimal multiuser codes when there are two users in the system, communicating simultaneously to a common base station. We assume that each user has n_t transmit antennas and there are n_r receive antennas at the receiving end. Further, we will assume a symmetric MAC channel [7], meaning the users transmit at same multiplexing gain r, or equivalently, both transmit at rate $R = r\log_2 \mathrm{SNR}$ in bits per channel use.

A. DMT for MIMO-MAC Channels

Considering a MIMO Rayleigh block fading channel, Tse *et al.* [7] showed that the codeword error probability of any such multiuser codes is lower bounded by

$$P_{\mathrm{cwe}}(\mathrm{SNR}) \dot{\geq} \max\left\{\mathrm{SNR}^{-d^*_{n_t,n_r}(r)}, \mathrm{SNR}^{-d^*_{2n_t,n_r}(2r)}\right\}, \tag{5}$$

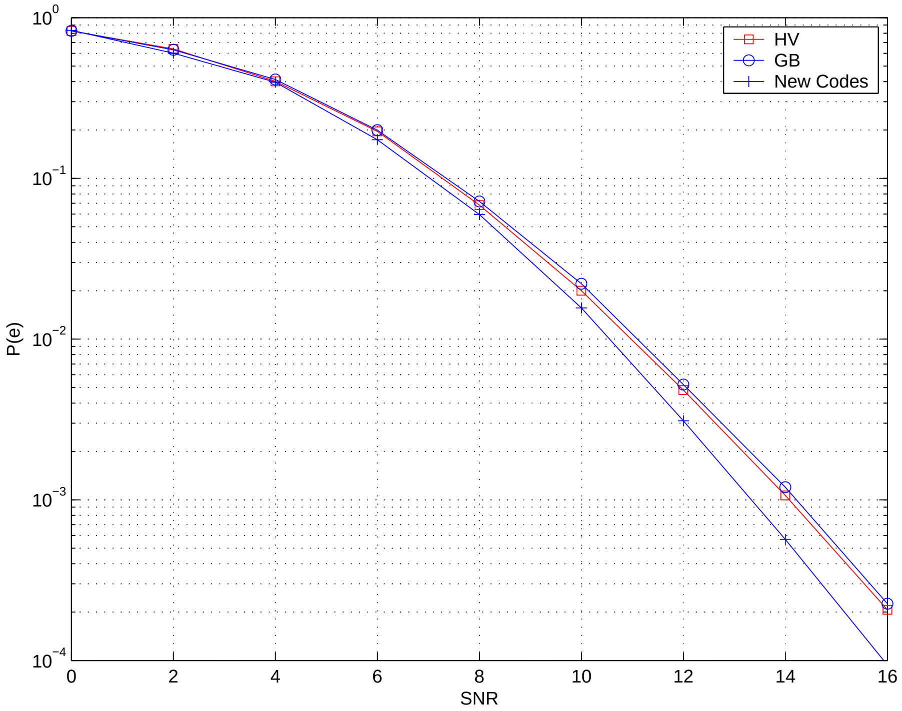

Fig. 1. The performance of the codes on 4-QAM received with 2 antennas.

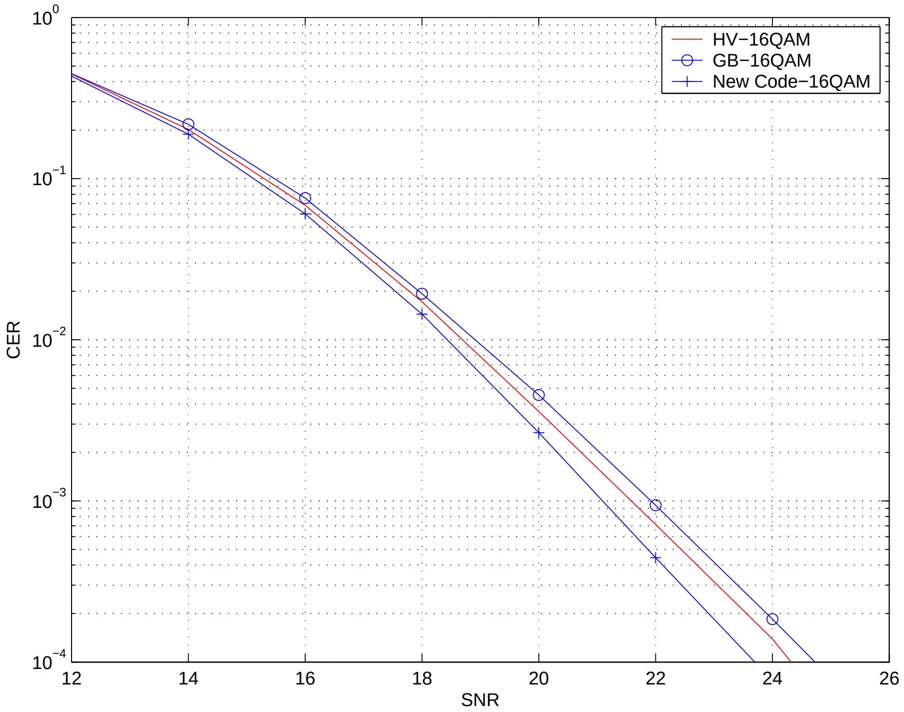

Fig. 2. The performance of the codes on 16-QAM received with 2 antennas.

where by $\dot{\geq}$ we mean the exponential inequality defined in [8], i.e. $f(\mathrm{SNR})\dot{\geq}g(\mathrm{SNR})$ if

$$\lim_{\mathrm{SNR}\to\infty}\frac{\log f(\mathrm{SNR})}{\log\mathrm{SNR}}\geq\lim_{\mathrm{SNR}\to\infty}\frac{\log g(\mathrm{SNR})}{\log\mathrm{SNR}}.$$

Notions of $\doteq$ and $\dot{\leq}$ are defined similarly.

The negative exponent $d^*_{n_t,n_r}(r)$ is the point-to-point DMT [8] for the case when there is only one user with n_t transmit antennas communicating at multiplexing gain r to the base station that has n_r receive antennas. $d^*_{n_t,n_r}(r)$ is a piecewise linear function connecting the points $(r,(n_t-r)(n_r-r))$ for $r=0,1,\cdots,\min\{n_t,n_r\}$. From this, in the two-user symmetric MIMO-MAC scenario, the maximal multiplexing gain can be achieved by the users is upper bounded by $r_{\max}=\min\{n_t,\frac{n_r}{2}\}$ since $d^*_{2n_t,n_r}(2r_{\max})=0$.

The terms $\mathrm{SNR}^{-d^*_{n_t,n_r}(r)}$ and $\mathrm{SNR}^{-d^*_{2n_t,n_r}(2r)}$ are respectively the probabilities when one or both users are in outage, i.e. the probabilities that the channel is not good enough to support the targeted rate. In particular, due to the behaviors of $d^*_{n_t,n_r}(r)$ and $d^*_{2n_t,n_r}(2r)$, Tse *et al.* showed that

$$\mathrm{SNR}^{-d^*_{n_t,n_r}(r)}\geq\mathrm{SNR}^{-d^*_{2n_t,n_r}(2r)},\ r\in\left[0,\min\{n_t,\frac{n_r}{3}\}\right].$$

That is, when $r\in\left[0,\min\{n_t,\frac{n_r}{3}\}\right]$, each user can achieve his/her best possible error performance as if the other user is not present in the channel. This is called the *single-user performance* regime. For $\min\{n_t,\frac{n_r}{3}\}\leq r\leq\min\{n_t,\frac{n_r}{2}\}$, the lower bound (5) is dominated by the second term, corresponding to the event of both users in outage. This is termed the *antenna pooling regime* [7]. These show a fundamental difference between single-user (or equivalently point-to-point) DMT and multiuser DMT.

By using independent Gaussian random codebooks for each user, the converse of (5) was proved by Tse *et al.* [7]. They partitioned the error events into two kinds, the kind when one of the two users is in error, denoted by $\mathscr{E}_1$, and the other kind when both users are in error, denoted by $\mathscr{E}_2$. They showed that when only one user is in error, the Gaussian random code is able to achieve an error performance with $\Pr\{\mathscr{E}_1\}\dot{\leq}\mathrm{SNR}^{-d^*_{n_t,n_r}(r)}$, and similarly $\Pr\{\mathscr{E}_2\}\dot{\leq}\mathrm{SNR}^{-d^*_{2n_t,n_r}(2r)}$ for the case when both users are in error. The above amounts to that given the multiplexing gain r, the maximal possible diversity gain can be achieved by any multiuser codes is $\min\{d^*_{n_t,n_r}(r),d^*_{2n_t,n_r}(2r)\}$. This is commonly referred to as the *optimal MAC-DMT*. Codes achieving this optimality are thus termed *MAC-DMT optimal* codes.

On the other hand, if deterministic codes were used; say code $\mathscr{S}_1$ for the first user and $\mathscr{S}_2$ for the second. Both codes consist of $(n_t\times T)$ code matrices for some T that corresponds to the channel coherence time, meaning the MIMO channel remains fixed during T symbol time. Further, the code matrices in $\mathscr{S}_1$ and $\mathscr{S}_2$ are required to satisfy the following power constraint:

$$E_{S_1\in\mathscr{S}_1}\|S_1\|_F^2\ \leq\ T\cdot\mathrm{SNR}\quad\text{and } E_{S_2\in\mathscr{S}_2}\|S_2\|_F^2\ \leq\ T\cdot\mathrm{SNR}. \tag{6}$$

By $\|A\|_F$ we mean the Frobenius norm of matrix A. Coronel *et al.* studied the optimal DMT performance of a selective fading MIMO multiple-access channel [13] and gave a sufficient criterion for designing MAC-DMT optimal multiuser codes. Noting that Rayleigh block fading channel can be regarded as a frequency selective fading channel with only one multipath, to our present interest, the criterion shown in [13] is equivalent to the following.

Theorem IV.1 ([13]). *Let $\mathscr{S}_1$ and $\mathscr{S}_2$ be defined as above with $T \geq 2n_t$. Then codes $\mathscr{S}_1$ and $\mathscr{S}_2$ achieve the optimal MAC-DMT if the following inequalities are all satisfied:*

$$\begin{aligned}\min_{S_1 \neq S_1' \in \mathscr{S}_1} \det\left((S_1 - S_1')(S_1 - S_1')^\dagger\right) &\ \dot{\geq}\ SNR^{n_t - r} \\ \min_{S_2 \neq S_2' \in \mathscr{S}_2} \det\left((S_2 - S_2')(S_2 - S_2')^\dagger\right) &\ \dot{\geq}\ SNR^{n_t - r} \\ \min_{S_1 \neq S_1' \in \mathscr{S}_1, S_2 \neq S_2' \in \mathscr{S}_2} \det\left(\Delta S \Delta S^\dagger\right) &\ \dot{\geq}\ SNR^{2n_t - 2r},\end{aligned}$$

where

$$\Delta S := \begin{bmatrix} S_1 - S_1' \\ S_2 - S_2' \end{bmatrix}$$

and where by $A^\dagger$ we mean the hermitian transpose of matrix A. □

We remark that the actual result in [13] was stated in a form different from the above. It includes the eigenvalues of channel matrix in all the above three conditions. However, by noting the equivalence between two constraint sets [22], [23] that is used to prove the property of approximate universal cyclic division algebra space-time codes [22], [23], [24], these results can be restated as in Theorem IV.1. For brevity, we do not elaborate on the details. However, if we set

$$\mathscr{C}_1 = \left\{ C_1 = \frac{1}{\kappa} S_1 \ : \ S_1 \in \mathscr{S}_1 \right\}$$

and similarly $\mathscr{C}_2 = \frac{1}{\kappa}\mathscr{S}_2$ with $\kappa^2 = \mathrm{SNR}^{1 - \frac{r}{n_t}}$, then the three criteria in Theorem IV.1 are equivalent to

$$\min_{C_1 \neq C_1' \in \mathscr{C}_1} \det\left((C_1 - C_1')(C_1 - C_1')^\dagger\right) \ \dot{\geq}\ 1 \tag{7}$$

$$\min_{C_2 \neq C_2' \in \mathscr{C}_2} \det\left((C_2 - C_2')(C_2 - C_2')^\dagger\right) \ \dot{\geq}\ 1 \tag{8}$$

$$\min_{C_1 \neq C_1' \in \mathscr{C}_1, C_2 \neq C_2' \in \mathscr{C}_2} \det\left(\Delta C \Delta C^\dagger\right) \ \dot{\geq}\ 1 \tag{9}$$

where $\Delta C = \frac{1}{\kappa}\Delta S$. Then we immediately recognize these are the well-known non-vanishing determinant (NVD) criteria [24], [23], [25], [26] for constructing point-to-point DMT optimal space-time codes. In other words, Theorem IV.1 is equivalent to the following. The proof can be regarded as an alternative proof to Theorem IV.1 in the flat fading case.

Theorem IV.2. *Let $\mathscr{C}_1$ and $\mathscr{C}_2$ be defined as above, and let the code $\mathscr{C}_1 \times \mathscr{C}_2$ be obtained by vertically concatenating the code matrices from $\mathscr{C}_1$ and $\mathscr{C}_2$. If $\mathscr{C}_1$, $\mathscr{C}_2$, and $\mathscr{C}_1 \times \mathscr{C}_2$ all satisfy NVD criterion, then the codes are MAC-DMT optimal.*

Proof: Similar to [7], we partition the error event into $\mathscr{E}_1$ and $\mathscr{E}_2$ that correspond respectively to the events when one or both users are in error. Then we have

$$\begin{aligned}\Pr\{\mathscr{E}_1\} &\leq P_{\text{cwe}}(\mathscr{C}_1) + P_{\text{cwe}}(\mathscr{C}_2) \ \dot{\leq}\ \mathrm{SNR}^{-d^*_{n_t,n_r}(r)} \\ \Pr\{\mathscr{E}_2\} &= P_{\text{cwe}}(\mathscr{C}_1 \times \mathscr{C}_2) \ \dot{\leq}\ \mathrm{SNR}^{-d^*_{2n_t,n_r}(2r)},\end{aligned}$$

where it follows from the fact that $\mathscr{C}_1$, $\mathscr{C}_2$, and $\mathscr{C}_1 \times \mathscr{C}_2$ are all DMT optimal in the point-to-point MIMO scenario. The readers are referred to [24] for the details. $P_{\text{cwe}}(\mathscr{C})$ denotes the codeword error probability of $\mathscr{C}$. □

Henceforth, we will refer to the criteria (7)-(9) as the *full NVD* condition. We note that as stated earlier, the full NVD condition is only sufficient for constructing MAC-DMT optimal codes, not necessary. In fact, we report the following negative result.[1]

Theorem IV.3. *When $n_t = 1$, i.e., each user with only one transmit antenna, there does not exist any multiuser codes that are full NVD.*

Proof: For ease of reading, the proof is relegated to the Appendix I. □

In a nutshell, the proof shows that while it is possible to construct DMT optimal codes $\mathscr{C}_1$ and $\mathscr{C}_2$ for user 1 and 2 respectively, as the existing cyclic-division algebra-based space-time codes [24] would do, it is impossible for the product code $\mathscr{C}_1 \times \mathscr{C}_2$ to be NVD. Any such product code would be ill-conditioned, i.e. having determinant extremely close to 0 at high SNR regime. It shows the nonexistence of codes satisfying the design criteria provided by Coronel *et al.* in [13]. Therefore, we may conclude that the full NVD condition is in general too strict to yield any MAC-DMT optimal codes. Another implication from the proof of Theorem IV.3 is the following. The full NVD condition can be met only if the two users cooperate in their transmission. Once without cooperation as it is in MIMO-MAC channel, the full NVD condition can never be met and the determinant must be vanishing.

However, we may relax the full NVD condition without affecting the DMT performance. To do so, we will use a different partition of error events. Let $\mathscr{E}_1$ denote again the event when one of the two users is in error. But let $\mathscr{E}_{2,1}$ (resp. $\mathscr{E}_{2,2}$) denote the error event when two users are in error and the error matrix is of rank n_t (resp. $2n_t$.) Clearly $\mathscr{E}_2$ is a disjoint union of $\mathscr{E}_{2,1}$ and $\mathscr{E}_{2,2}$. Now the codes $\mathscr{C}_1$ and $\mathscr{C}_2$ are MAC-DMT optimal if the following holds.

Theorem IV.4. *Let $\mathscr{C}_1$ and $\mathscr{C}_2$ be defined as above. Then they are MAC-DMT optimal if the error events have probabilities upper bounded by*

$$\begin{aligned}
\Pr\{\mathscr{E}_1\} &\dot{\leq} SNR^{-d^*_{n_t,n_r}(r)}, \\
\Pr\{\mathscr{E}_{2,1}\} &\dot{\leq} SNR^{-d^*_{n_t,n_r}(r)}, \\
\Pr\{\mathscr{E}_{2,2}\} &\dot{\leq} SNR^{-d^*_{2n_t,n_r}(2r)}.
\end{aligned}$$

□

The rationale behind the above theorem is the observation that in the single-user performance regime, the error probability $\text{SNR}^{-d^*_{2n_t,n_r}(2r)}$ is not dominant, hence we could relax the condition such that event $\mathscr{E}_{2,1}$ has larger probability $\text{SNR}^{-d^*_{n_t,n_r}(r)}$ than the actual outage probability $\text{SNR}^{-d^*_{2n_t,n_r}(2r)}$. This will not

[1] A more general result of the nonexistence of full NVD multiuser codes that satisfy the criteria given by Coronel *et al.* [13] for arbitrary number of transmit antennas and for arbitrary number of users has been proved by the authors, but it will be treated in a separate paper [14].

affect the overall DMT performance. Compared with the full NVD condition required in Theorems IV.1 and IV.2, Theorem IV.4 relaxes greatly the code design criterion. Specifically, the full NVD condition requires that whenever $C_1 \neq C_1' \in \mathscr{C}_1$ and $C_2 \neq C_2' \in \mathscr{C}_2$, the matrix ΔC must be nonsingular and be NVD, i.e. having determinant $\det(\Delta C \Delta C^\dagger) \geq 1$. This has been shown to be impossible by Theorem IV.3. On the other hand, Theorem IV.4 says that the difference matrix ΔC can be singular, and the only condition is that should it happen, the resulting error performance cannot be worse than $\mathrm{SNR}^{-d^*_{n_t,n_r}(r)}$, in order to maintain the MAC-DMT optimality. In [13], event $\mathscr{E}_{2,1}$ was required to have probability absolutely zero, which is too strict and forbids the existence of MAC-DMT optimal codes.

B. *Construction of MAC-DMT Optimal Codes*

In this section, we will provide a systematic construction of multiuser codes for the two-user case. The proposed codes will not meet the full NVD criterion as such codes do not exist. In the next section we will analyze the DMT performance of these newly proposed codes and show that they actually achieve the relaxed criteria given in Theorem IV.4.

Let $F = \mathbf{Q}(i)$ be the base number field. The proposed construction calls for two additional number fields $L = F(\theta)$ and $K = F(\eta)$ that are cyclic Galois extension of F with $[L:F] = n_t$ and $[K:F] = 2$. We require further that $L \cap K = F$. Let $\mathrm{Gal}(L/F) = \langle \sigma \rangle$ and $\mathrm{Gal}(K/F) = \langle \tau \rangle$, and let $E = LK = F(\theta, \eta)$ be the compositum of the fields L and K. The relation between these field extensions is shown in Fig. 3.

$$\begin{tikzcd}
& E = F(\theta,\eta) \arrow[dl, dash, "\langle\tau\rangle"'] \arrow[dr, dash, "\langle\sigma\rangle"] & \\
L = F(\theta) \arrow[dr, dash, "\langle\sigma\rangle"'] & & K = F(\eta) \arrow[dl, dash, "\langle\tau\rangle"] \\
& F = \mathbf{Q}(i) &
\end{tikzcd}$$

Fig. 3. Field extensions required by the proposed code constructions.

Clearly, L/F is cyclic Galois; so is E/K. Moreover, we have $\mathrm{Gal}(E/K) = \langle \sigma \rangle$. Hence there exists some suitable non-norm element $\gamma \in \mathscr{O}_F$ such that

$$\mathfrak{D} = (E/K, \sigma, \gamma) = E \oplus uE \oplus \cdots \oplus u^{n_t - 1}E$$

is a division algebra, where by $\mathscr{O}_F$ we mean the ring of algebraic integers in F and u is an indeterminate satisfying $u^{n_t} = \gamma$ and $xu = u\sigma(x)$ for every $x \in E$. Similarly as in Section II, let again $\psi : \mathfrak{D} \to M_{n_t}(E)$ be the left-regular map that represents every element $x = \sum_{i=0}^{n_t-1} u^i x_i \in \mathfrak{D}$, $x_i \in E$, as an $n_t \times n_t$ matrix given by

$$\psi(x) \;:=\; \begin{pmatrix} x_0 & \gamma\sigma(x_{n_t-1}) & \cdots & \gamma\sigma^{n_t-1}(x_1) \\ x_1 & \sigma(x_0) & \cdots & \gamma\sigma^{n_t-1}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ x_{n_t-1} & \sigma(x_{n_t-2}) & \cdots & \sigma^{n_t-1}(x_0) \end{pmatrix}. \tag{10}$$

According to Definition II.7 and Proposition II.4 $\det(\psi(x)) \in K$ for every $x \in \mathfrak{D}$, and hence clearly

$$nr_{K/F}\left(\det(\psi(x))\right) = \det(\psi(x))\tau\left(\det(\psi(x))\right) \in F \tag{11}$$

where $nr_{K/F}(a)$ is the algebraic norm of a from K to F. Note that when the element x is taken from the natural order $\mathscr{O}_{\mathfrak{D}} := \mathscr{O}_E \oplus \cdots \oplus u^{n_t-1}\mathscr{O}_E$, it can be further shown that

$$nr_{K/F}\left(\det(\psi(x))\right) = \det(\psi(x))\tau\left(\det(\psi(x))\right) \in \mathscr{O}_F \tag{12}$$

and $\mathscr{O}_F = \mathbf{Z}[i]$. It in turn implies that the absolute $\left|nr_{K/F}\left(\det(\psi(x))\right)\right|$ is bounded from below by 1 whenever $0 \neq x \in \mathscr{O}_{\mathfrak{D}}$. This property is termed *generalized non-vanishing determinant condition* in [23] (also cf. Definition III.2) and is required in constructing the DMT optimal multi-block space-time codes.

Having said the above, the proposed construction is the following. Given the multiplexing gain r, let

$$\mathscr{A}(\text{SNR}) = \left\{a + bi : -\text{SNR}^{\frac{r}{2n_t}} \leq a, b \leq \text{SNR}^{\frac{r}{2n_t}}, \quad a, b \text{ odd}\right\} \tag{13}$$

and let $\{e_0, \cdots, e_{2n_t-1}\}$ be an integral basis of E/F. Given $\mathscr{A}(\text{SNR})$ we define the information set

$$\mathfrak{A}(\text{SNR}) \;=\; \left\{\sum_{i=0}^{n_t-1} u^i \sum_{j=0}^{2n_t-1} a_{i,j}e_i \;:\; a_{i,j} \in \mathscr{A}(\text{SNR})\right\}. \tag{14}$$

It is clear that $\mathfrak{A}(\text{SNR}) \subset \mathscr{O}_{\mathfrak{D}}$.

If the first user wishes to transmit information $x \in \mathfrak{A}(\text{SNR})$, the transmitter actually sends in $2n_t$ channel uses the $(n_t \times 2n_t)$ code matrix

$$S_x \;=\; \kappa\begin{pmatrix} \psi(x) & \tau\left(\psi(x)\right) \end{pmatrix}, \tag{15}$$

where κ is a constant given by

$$\kappa^2 \;\doteq\; \text{SNR}^{1-\frac{r}{n_t}} \tag{16}$$

and is set such that $E\,\|S_x\|_F^2 = 2n_t \cdot \text{SNR}$.

On the other hand, if the second user wishes to transmit information $y \in \mathfrak{A}(\text{SNR})$, the resulting code matrix associated with y is

$$S_y \;=\; \kappa\begin{pmatrix} \psi(y) & -\tau\left(\psi(y)\right) \end{pmatrix}. \tag{17}$$

With regard to the channel model, given the transmitted code matrices S_x and S_y from the first and the second users, respectively, let H_1 and H_2 be respectively the $(n_r \times n_t)$ channel matrices associated with the first and the second users. The overall received signal matrix R_o is given by

$$R_o \;=\; H_1S_x + H_2S_y + W = \kappa\begin{pmatrix} H_1 & H_2 \end{pmatrix}\begin{pmatrix} X & \tau(X) \\ Y & -\tau(Y) \end{pmatrix} + W \tag{18}$$

where $X := \psi(x)$, $Y := \psi(y)$, and where W is the $(n_r \times 2n_t)$ noise matrix whose entries are i.i.d. $\mathbf{C}\mathscr{N}(0,1)$ random variables. Therefore, our proposed multiuser code may be described as follows

$$\mathscr{S} = \left\{ \kappa \begin{pmatrix} X & \tau(X) \\ Y & -\tau(Y) \end{pmatrix} : \begin{array}{l} X = \psi(x), Y = \psi(y), \\ x, y \in \mathfrak{A}(\mathrm{SNR}) \end{array} \right\}. \tag{19}$$

For every code matrix $S \in \mathscr{S}$, the upper half submatrix corresponds to the information sent by the first user and the lower half comes from the second user. Clearly the two submatrices are coded independently, and there is no cooperation between these two users.

As κ is a normalizing constant for power constraint, below we will pay our attention only to the set of unnormalized code matrices, i.e.

$$\mathscr{C} = \left\{ \begin{pmatrix} X & \tau(X) \\ Y & -\tau(Y) \end{pmatrix} : \begin{array}{l} X = \psi(x), Y = \psi(y), \\ x, y \in \mathfrak{A}(\mathrm{SNR}) \end{array} \right\}. \tag{20}$$

First, we show that every code matrix $C \in \mathscr{C}$ has determinant in $\mathbf{Z}[i]$.

Lemma IV.5. *Let $\mathscr{C}$ be defined as above; then for every $C \in \mathscr{C}$,* $\det(C) \in \mathbf{Z}[i]$.

Proof: Clearly, the entries of C lie in $\mathscr{O}_E$, the ring of algebraic integers in E; hence $\det(C) \in \mathscr{O}_E$. It suffices to show that the determinant is fixed by the automorphisms τ and σ. To this end, given any $C \in \mathscr{C}$, we simply check

$$\begin{aligned} \tau(\det(C)) &= \det \begin{pmatrix} \tau(X) & X \\ \tau(Y) & -Y \end{pmatrix} = (-1)^{n_t} \det \begin{pmatrix} X & \tau(X) \\ -Y & \tau(Y) \end{pmatrix} \\ &= (-1)^{n_t} \det \left(\begin{pmatrix} I_{n_t} & \\ & -I_{n_t} \end{pmatrix} \begin{pmatrix} X & \tau(X) \\ Y & -\tau(Y) \end{pmatrix} \right) \\ &= (-1)^{2n_t} \det(C) = \det(C) \end{aligned}$$

and

$$\begin{aligned} \sigma(\det(C)) &= \det \begin{pmatrix} Z^{-1}XZ & \tau(Z^{-1}XZ) \\ Z^{-1}YZ & -\tau(Z^{-1}YZ) \end{pmatrix} = \det \begin{pmatrix} Z^{-1}XZ & Z^{-1}\tau(X)Z \\ Z^{-1}YZ & -Z^{-1}\tau(Y)Z \end{pmatrix} \\ &= \det \left(\begin{pmatrix} Z^{-1} & \\ & Z^{-1} \end{pmatrix} \begin{pmatrix} X & \tau(X) \\ Y & -\tau(Y) \end{pmatrix} \begin{pmatrix} Z & \\ & Z \end{pmatrix} \right) = \det(C) \end{aligned}$$

where $Z := \psi(u)$ and where we have used the fact that $\tau(Z) = Z$ as $\gamma \in \mathscr{O}_F$. Overall, these show $\det(C) \in \mathbf{Z}[i]$. □

While the above lemma shows that the determinant of the matrix C lies in $\mathbf{Z}[i]$, it does not necessarily mean that the code satisfies the NVD property. For example, if $\tau : \eta \to -\eta$, then setting $y = \eta x \in \mathfrak{A}(\mathrm{SNR})$ makes the resulting code matrix C singular as the lower half can be obtained by multiplying from the left the upper half by matrix $\psi(\eta)$. In particular, whether the code matrix C is singular or not, is completely characterized by the following lemma.

Lemma IV.6. *Given*

$$C=\begin{pmatrix} X & \tau(X) \\ Y & -\tau(Y) \end{pmatrix} \in \mathscr{C}$$

with $X=\psi(x)$ and $Y=\psi(y)$, $x,y\in\mathfrak{A}(SNR)$, if $x\neq 0$, then

$$rank(C) \;=\; \begin{cases} n_t, & \text{if } yx^{-1}+\tau(yx^{-1})=0 \\ 2n_t, & \text{otherwise.} \end{cases} \tag{21}$$

Moreover, if $\tau:\eta\to-\eta$ then $rank(C)=n_t$ if and only if

$$yx^{-1}\in\bigoplus_{i=0}^{n_t-1}u^i\eta L:=\mathfrak{L}. \tag{22}$$

Proof: To find out the rank of matrix C, we follow the conventional Gaussian eliminant procedure with elementary row operations. In particular, we remark that such operations would be easier to carry out if we change our focus to the matrix

$$\tilde{C} \;=\; \begin{pmatrix} x & \tau(x) \\ y & -\tau(y) \end{pmatrix} \in M_2(\mathfrak{D}).$$

This is because elementary row operations in $M_2(\mathfrak{D})$ correspond exactly to block elementary row operations in C. Specifically, we mean following

$$\psi\left(\begin{pmatrix} p & q \end{pmatrix}\tilde{C}\right)=\begin{pmatrix} \psi(p) & \psi(q) \end{pmatrix}C.$$

Thus, if $x\neq 0$ by assumption we see that $\text{rank}(\psi(x))=n_t$ as $\mathfrak{D}$ is a division algebra, and secondly that there must exist $p\in\mathfrak{D}$ such that $y=px$ since $yx^{-1}\in\mathfrak{D}$. Then we can rewrite $\tilde{C}$ as

$$\tilde{C} \;=\; \begin{pmatrix} x & \tau(x) \\ px & -\tau(p)\tau(x) \end{pmatrix}.$$

Multiplying from the left the first row of $\tilde{C}$ by $-p$ and adding to the second row yields

$$\begin{pmatrix} x & \tau(x) \\ 0 & -\left(\tau(p)+p\right)\tau(x) \end{pmatrix}.$$

It is clear that $\tilde{C}$ is left- and right- invertible in $M_{n_t}(\mathfrak{D})$ if and only if $\tau(p)+p\neq 0$. In other words, C is singular if and only if $yx^{-1}+\tau(yx^{-1})=0$.

To prove the second claim, we first note that $\{1,\theta,\cdots,\theta^{n_t-1}\}$ is a basis of L/F and similarly $\{1,\eta\}$ a basis for K/F. $p=yx^{-1}$ can be uniquely represented as

$$p \;=\; \sum_{i=0}^{n_t-1}u^i\sum_{j=0}^{n_t-1}p_{1,i,j}\theta^j+\sum_{i=0}^{n_t-1}u^i\eta\sum_{j=0}^{n_t-1}p_{2,i,j}\theta^j$$

for some $p_{1,i,j},p_{2,i,j}\in F$. Hence

$$\tau(p)=\sum_{i=0}^{n_t-1}u^i\sum_{j=0}^{n_t-1}p_{1,i,j}\theta^j-\sum_{i=0}^{n_t-1}u^i\eta\sum_{j=0}^{n_t-1}p_{2,i,j}\theta^j.$$

Now we see $p=-\tau(p)$ if and only if $p_{1,i,j}=0$ for all i and j. This proves the claim. □

Remark IV.1. The above lemma shows that the proposed construction does not satisfy the full NVD criterion. This is not surprising as already pointed out in Theorem IV.3 that codes satisfying full NVD criterion do not exist. Yet, as suggested by the reviewers, it is sometimes interesting to see how often the code violates the full NVD criterion. That is, we are interested in knowing $\Pr\{p+\tau(p)=0\}$. Although such probability depends closely upon the underlying set of base alphabet $\mathscr{A}(\mathrm{SNR})$, we can argue heuristically to show such probability is extremely small. Furthermore, our estimate of $\Pr\{p+\tau(p)=0\}$ will be asymptotically tight at high SNR regime, i.e. when the transmission rate R (in bits per channel use) gets larger and larger.

To see the above, let us fix x, the symbol sent by the first user and consider all possible choices of y sent by the second user. Clearly, as $p=yx^{-1}\in\mathfrak{D}$ we have $p \ = \ p_0+up_1+\cdots+u^{n_t-1}p_{n_t-1}$ with $p_i\in E$. Define

$$\mathscr{P} \ := \ \left\{p=yx^{-1}: y\in\mathfrak{A}(\mathrm{SNR}), p+\tau(p)=0\right\}.$$

Note that from (22) we have

$$|\mathscr{P}|=\left|\left\{p=yx^{-1}: y\in\mathfrak{A}(\mathrm{SNR}), p\in\mathfrak{L}\right\}\right| \ \dot{\leq}\ |\{z\in\mathfrak{A}(\mathrm{SNR}) \ : \ z\in\mathfrak{L}\}| \ = \ |\mathscr{A}(\mathrm{SNR})|^{n_t^2}.$$

The inequality $\dot{\leq}$ is because of the following. Given any $p=\sum_{i=0}^{n_t-1}u^ip_i$ with $p+\tau(p)=0$, the element

$$y \ = \ px \ = \ \sum_{i=0}^{n_t-1}u^i\sum_{j=0}^{2n_t-1}y_{i,j}e_j$$

might not be in $\mathfrak{A}(\mathrm{SNR})$, since

1) the element $y_{i,j}$ might not be a Gaussian integer, and
2) $y_{i,j}$ might not be in $\mathscr{A}(\mathrm{SNR})$, especially when $\mathscr{A}(\mathrm{SNR})$ is of small size.

Thus, the above estimate of $|\mathscr{P}|$ is generally loose for small $\mathscr{A}(\mathrm{SNR})$. However, when $\mathscr{A}(\mathrm{SNR})$ becomes larger, px is likely to be in $\mathfrak{A}(\mathrm{SNR})$ and the proposed estimate becomes more accurate. Overall, as $|\mathfrak{A}(\mathrm{SNR})|=\mathrm{SNR}^{2n_tr}$ we see

$$\Pr\{p+\tau(p)=0\} \ \dot{\leq}\ \frac{|\mathscr{P}|}{|\mathfrak{A}(\mathrm{SNR})|} \doteq \frac{1}{\sqrt{|\mathfrak{A}(\mathrm{SNR})|}} \doteq \mathrm{SNR}^{-n_tr}. \tag{23}$$

When $n_t=2$, we numerically simulated the probability $\Pr\{p+\tau(p)=0\}$ at different rates.

- At $R=4$ and $\mathscr{A}(\mathrm{SNR})$ being QPSK, the probability $\Pr\{p+\tau(p)=0\}\approx 5.15\times 10^{-5}$, while (23) gives $4^{-4}\approx 4\times 10^{-3}$.
- At $R=6$ and $\mathscr{A}(\mathrm{SNR})$ being 8QAM, we get $\Pr\{p+\tau(p)=0\}\approx 1.104\times 10^{-8}$, while (23) gives $8^{-4}\approx 2\times 10^{-4}$.
- At $R=8$ and $\mathscr{A}(\mathrm{SNR})$ being 16QAM, we report $\Pr\{p+\tau(p)=0\}\approx 1.194\times 10^{-9}$, while (23) gives $16^{-4}\approx 10^{-5}$.

Thus we see in general for high transmission rate, $\Pr\{p+\tau(p)=0\}$ is extremely close to 0, and the difference matrix ΔC is of full rank with probability close to 1. Furthermore, from the simulations above we see that at small size of $\mathscr{A}(\mathrm{SNR})$, the probability $\Pr\{p+\tau(p)=0\}$ behaves more like

$$\Pr\{p+\tau(p)=0\} \ \approx \ \frac{|\mathscr{A}(\mathrm{SNR})|}{|\mathfrak{A}(\mathrm{SNR})|} \ = \ |\mathscr{A}(\mathrm{SNR})|^{-(2n_t^2-1)}$$

since not all $y = px$ belong to $\mathfrak{A}(\text{SNR})$ for a fixed x and a random p with $p + \tau(p) = 0$. □

Armed with the two above lemmas, we are now ready to show that the proposed code $\mathscr{S}$ is MAC-DMT optimal. The proof will be given in the next subsection.

Theorem IV.7. *Given the multiplexing gain r, the proposed code $\mathscr{S}$ achieves over quasi-static Rayleigh fading channel with coherence time $T \geq 2n_t$ the DMT*

$$d(r) = \begin{cases} d^*_{n_t,n_r}(r), \text{ if } r \leq \min\left\{n_t, \frac{n_r}{3}\right\} \\ d^*_{2n_t,n_r}(2r), \text{ if } r \in \left(\min\left\{n_t, \frac{n_r}{3}\right\}, \min\left\{n_t, \frac{n_r}{2}\right\}\right) \end{cases} \tag{24}$$

meaning that $\mathscr{S}$ is MAC-DMT optimal. □

C. Proof of Theorem IV.7

For any $S \neq S' \in \mathscr{S}$ with

$$S = \kappa \begin{pmatrix} \psi(x) & \tau(\psi(x)) \\ \psi(y) & -\tau(\psi(y)) \end{pmatrix}, \quad S' = \kappa \begin{pmatrix} \psi(x') & \tau(\psi(x')) \\ \psi(y') & -\tau(\psi(y')) \end{pmatrix},$$

define $dx := x - x'$ and $dy := y - y'$. Hence

$$\Delta S = S - S' = \kappa \begin{pmatrix} \psi(dx) & \tau(\psi(dx)) \\ \psi(dy) & -\tau(\psi(dy)) \end{pmatrix}. \tag{25}$$

Following Theorem IV.3, we will be considering the following error events:

1) Event $\mathscr{E}_1$ corresponds to the case when either user one or user two is in error, but not both. This means that the difference matrix ΔS of (25) has either $dx = 0$ or $dy = 0$.
2) Error event $\mathscr{E}_{2,1}$ concerns the case when both users are in error, but the overall error matrix ΔS is not of full rank $2n_t$. That is, we have both dx and dy being nonzero, but the error matrix ΔS is only of rank n_t and $dy(dx)^{-1} + \tau(dy(dx)^{-1}) = 0$.
3) Error event $\mathscr{E}_{2,2}$ is the case when both users are in error and the error matrix ΔS is of full rank $2n_t$.

Clearly, whenever a decoding error occurs, the error event $\mathscr{E}$ is a union of the above three error events, namely, we have

$$\mathscr{E} = \mathscr{E}_1 \cup \mathscr{E}_{2,1} \cup \mathscr{E}_{2,2}$$

and the corresponding error probability achieved by $\mathscr{S}$ is

$$P_{\text{cwe}}(\text{SNR}) = \Pr\{\mathscr{E}\} \leq \Pr\{\mathscr{E}_1\} + \Pr\{\mathscr{E}_{2,1}\} + \Pr\{\mathscr{E}_{2,2}\}.$$

Thus, in the remaining of this section we will show

$$\begin{aligned} \Pr\{\mathscr{E}_1\} &\dot{\leq} \text{SNR}^{-d^*_{n_t,n_r}(r)}, \\ \Pr\{\mathscr{E}_{2,1}\} &\dot{\leq} \text{SNR}^{-d^*_{n_t,n_r}(r)}, \\ \Pr\{\mathscr{E}_{2,2}\} &\dot{\leq} \text{SNR}^{-d^*_{2n_t,n_r}(2r)}. \end{aligned}$$

a) Error Event $\mathscr{E}_1$: We first focus on analyzing the error event $\mathscr{E}_1$ that corresponds to the case when either user one or user two is in error, but not both. Given the channel matrices H_1 and H_2 we define the squared Euclidean distance between S and S' as

$$d_E^2(S,S') \ := \ \|H\Delta S\|_F^2 \tag{26}$$

where $H = [H_1\ H_2]$. Due to the structure of $\mathscr{S}$, we can without loss of generality assume that $dx \neq 0$ but $dy = 0$. The other case of $dx = 0$, $dy \neq 0$ can be analyzed in a similar fashion. Thus in this case we have

$$d_E^2(S,S') = \|H_1\psi(dx)\|_F^2 + \|H_1\tau(\psi(dx))\|_F^2\,. \tag{27}$$

To obtain a lower bound on $d_E^2(S,S')$, let $\lambda_{1,1} \geq \cdots \geq \lambda_{1,m}$ be the set of ordered nonzero eigenvalues of $H_1H_1^\dagger$ where $m = \min\{n_t, n_r\}$ and let $\ell_{1,1} \leq \cdots \leq \ell_{1,n_t}$ and $\ell_{2,1} \leq \cdots \leq \ell_{2,n_t}$ be the ordered nonzero eigenvalues of $\psi(dx)\psi(dx)^\dagger$ and $\tau(\psi(dx))\tau(\psi(dx))^\dagger$, respectively. Using the mismatch eigenvalue bound [27], [24], [23] we see $d_E^2(S,S')$ is lower bounded by

$$d_E^2(S,S') \ \geq \ \kappa^2 \sum_{i=1}^{m} \lambda_{1,i}\,(\ell_{1,n_t-m+i} + \ell_{2,n_t-m+i})\,. \tag{28}$$

Note that

$$\prod_{i=1}^{n_t}\prod_{j=1}^{2} \ell_{j,i} = \left|nr_{K/F}\left(\det(\psi(dx))\right)\right|^2 \geq 1. \tag{29}$$

Repeatedly using the arithmetic mean-geometric mean inequality and (29) along the same lines as in [24], [23], given k, $k = 1,2,\cdots,m$, it can be shown that

$$\begin{aligned} d_E^2(S,S') \ &\dot{\geq} \ \kappa^2 \left[\prod_{i=m-k+1}^{m} \lambda_{1,i}\right]^{\frac{1}{k}} \left[\|\psi(dx)\|_F^2 + \|\tau(\psi(dx))\|_F^2\right]^{-\frac{n_t-k}{k}} \\ &\dot{\geq} \ \mathrm{SNR}^{1-\frac{r}{n_t}} \left[\prod_{i=m-k+1}^{m} \lambda_{1,i}\right]^{\frac{1}{k}} \mathrm{SNR}^{-\frac{r}{n_t}\frac{n_t-k}{k}}. \end{aligned}$$

Setting $\lambda_{1,i} \ = \ \mathrm{SNR}^{-\alpha_{1,i}}$ gives

$$d_E^2(S,S') \ \dot{\geq} \ \mathrm{SNR}^{\delta_{1,k}(\underline{\alpha}_1)} \tag{30}$$

where $\underline{\alpha}_1 = [\alpha_{1,1}\cdots\alpha_{1,m}]^t$ and

$$\delta_{1,k}(\underline{\alpha}_1) := \frac{1}{k}\left[\sum_{i=m-k+1}^{m}(1-\alpha_{1,i})\right] - \frac{r}{k}. \tag{31}$$

Following the sphere bound argument as in [24], the probability of event $\mathscr{E}_1$ given the channel matrices H_1 and H_2 can be upper bounded by

$$\Pr\{\mathscr{E}_1|H_1,H_2\} \leq \Pr\left\{\|W\|_F^2 \ \geq \ \frac{d_E^2(S,S')}{4}\right\} = \exp\left(-\frac{d_E^2(S,S')}{4}\right)\sum_{j=0}^{2n_rn_t-1}\frac{(d_E^2(S,S'))^j}{j!}.$$

As $d_E^2(S,S')\dot{\geq}\mathrm{SNR}^{\delta_{1,k}(\underline{\alpha}_1)}$ for all k, we see from the above that $\Pr\{\mathscr{E}_1|H_1,H_2\} \doteq 0$ if there exists k such that $\delta_{1,k}(\underline{\alpha}_1) > 0$. Since $\Pr\{\mathscr{E}_1|H_1,H_2\} \leq 1$, it follows that

$$\Pr\{\mathscr{E}_1\} = E_{H_1,H_2}\Pr\{\mathscr{E}_1|H_1,H_2\} \leq 2\Pr\left\{\underline{\alpha}_1 : \delta_{1,k}(\underline{\alpha}_1) \leq 0, \ \text{all } k\right\},$$

where the extra factor of 2 shown above is due to the inclusion of the other case when user two is in error which has the same probability as the present case. Clearly, in terms of diversity analysis one can safely neglect this factor of 2.

Arguing similarly as [23], [22] it can be shown that

$$\left\{\underline{\alpha}_1 : \delta_{1,k}(\underline{\alpha}_1) \leq 0, k = 1, \cdots, m\right\} = \left\{\underline{\alpha}_1 : \sum_{i=1}^{m} (1-\alpha_{1,i})^+ \ \leq \ r\right\} \tag{32}$$

where $(x)^+ := \max\{x, 0\}$. Now we see

$$\begin{aligned} \Pr\{\mathscr{E}_1\} &\leq \Pr\left\{\underline{\alpha}_1 : \sum_{i=1}^{m} (1-\alpha_{1,i})^+ \ \leq \ r\right\} = \Pr\left\{\log\det\left(I_{n_r} + \mathrm{SNR} H_1 H_1^\dagger\right) \leq r \log \mathrm{SNR}\right\} \\ &\doteq \mathrm{SNR}^{-d^*_{n_t,n_r}(r)}, \end{aligned}$$

where the last exponential equality follows from [8].

b) Error Event $\mathscr{E}_{2,2}$: For simplicity, we will first analyze the event $\mathscr{E}_{2,2}$, and leave the most tedious event $\mathscr{E}_{2,1}$ to the last. Recall that $\mathscr{E}_{2,2}$ is the event when both users are in error, and the error matrix ΔS is of full rank $2n_t$. In other words, we have in (25) that $dx, dy \neq 0$ and $dy\,(dx)^{-1} + \tau\left(dy\,(dx)^{-1}\right) \neq 0$. Lemmas IV.5 and IV.6 then imply the matrix

$$\Delta C = \begin{pmatrix} \psi(dx) & \tau(\psi(dx)) \\ \psi(dy) & -\tau(\psi(dy)) \end{pmatrix} \tag{33}$$

must have full rank $2n_t$ and $1 \leq |\det(\Delta C)| \in \mathbf{Z}$. Let $\ell_1 \leq \ell_2 \leq \cdots \leq \ell_{2n_t}$ be the ordered eigenvalues of $\Delta C \Delta C^\dagger$, and let $\lambda_{2,1} \geq \cdots \geq \lambda_{2,m'}$ be the ordered nonzero eigenvalues of $HH^\dagger$ with $H = [H_1 \ H_2]$ and $m' = \min\{2n_t, n_t\}$.

Following arguments similar to $\mathscr{E}_1$, the squared Euclidean distance $d_E^2(S,S')$ for the pair (S,S') falling in the category of $\mathscr{E}_{2,2}$ is lower bounded by

$$\begin{aligned} d_E^2(S,S') &\geq \kappa^2 \sum_{i=m'-k+1}^{m'} \lambda_{2,i}\ell_{2n_t-m'+i} \geq \kappa^2 \left[\prod_{i=m'-k+1}^{m'} \lambda_{2,i}\right]^{\frac{1}{k}} \left[\prod_{i=1}^{2n_t-k} \ell_i\right]^{-\frac{1}{k}} \\ &\dot{\geq} \kappa^2 \left[\prod_{i=m'-k+1}^{m'} \lambda_{2,i}\right]^{\frac{1}{k}} \left[\sum_{i=1}^{2n_t-k} \ell_i\right]^{-\frac{2n_t-k}{k}} \dot{\geq} \mathrm{SNR}^{1-\frac{r}{n_t}} \left[\prod_{i=m'-k+1}^{m'} \lambda_{1,i}\right]^{\frac{1}{k}} \mathrm{SNR}^{-\frac{r}{n_t}\frac{2n_t-k}{k}} \\ &:= \mathrm{SNR}^{\delta_{2,k}(\underline{\alpha}_2)}, \end{aligned}$$

for $k = 1, 2, \cdots, m'$, where $\lambda_{2,i} = \mathrm{SNR}^{-\alpha_{2,i}}$ and

$$\delta_{2,k}\left(\underline{\alpha}_2\right) := \frac{1}{k}\left[\sum_{i=m'-k+1}^{m'} (1-\alpha_{2,i})\right] - \frac{2r}{k}. \tag{34}$$

Again along the same lines as in the previous case we can show that

$$\begin{aligned} \Pr\left\{\mathscr{E}_{2,2}\right\} &\dot{\leq} \Pr\left\{\underline{\alpha}_2 : \delta_{2,k}(\underline{\alpha}_2) \leq 0, \text{ all } k\right\} = \Pr\left\{\underline{\alpha}_2 : \sum_{i=1}^{m'} (1-\alpha_{2,i})^+ \ \leq \ 2r\right\} \\ &= \Pr\left\{\log\det\left(I_{n_r} + \mathrm{SNR} HH^\dagger\right) \leq 2r \log \mathrm{SNR}\right\} \doteq \mathrm{SNR}^{-d^*_{2n_t,n_r}(2r)}, \end{aligned}$$

proving that the code $\mathscr{S}$ satisfies the third condition required in Theorem IV.4.

c) Error Event $\mathscr{E}_{2,1}$: Finally we are left with the last type of error event, the event $\mathscr{E}_{2,1}$ occurring when both users are in error, but the error matrix does not have full rank. In other words, it is the case when $dx, dy \neq 0$, $p = dy\,(dx)^{-1}$ and $p + \tau(p) = 0$ in (25). From the proof of Lemma IV.6 these conditions mean

$$\psi(dy) = P\psi(dx) \text{ and } \tau(-\psi(dy)) = P\tau(\psi(dx)),$$

where $P = \psi(p)$ is nonsingular in $M_{n_t}(E)$ and where we have used the fact that $P + \tau(P) = \mathbf{0}$. Thus, the squared Euclidean distance $d_E^2(S,S')$ for the pair (S,S') in this category can be rewritten as

$$\begin{aligned} d_E^2(S,S') &= \|H_1\psi(dx) + H_2P\psi(dx)\|_F^2 + \|H_1\tau(\psi(dx)) + H_2P\tau(\psi(dx))\|_F^2 \\ &= \left\|H_{3p}\psi(dx)\right\|_F^2 + \left\|H_{3p}\tau(\psi(dx))\right\|_F^2, \end{aligned} \tag{35}$$

where $H_{3p} := H_1 + H_2P$. We keep the p in the subscript of H_{3p} to indicate that H_{3p} is a function of the ratio p for different pairs of (dx,dy) with the required properties. For any p, the matrix H_{3p} is of full rank with probability one, and we can let $\lambda_{3p,1} \geq \cdots \geq \lambda_{3p,m}$ be the ordered nonzero eigenvalues of $H_{3p}H_{3p}^\dagger$ with $m = \min\{n_t, n_t\}$. Note that (35) is the same as (27) except that the channel matrix H_1 is replaced by H_{3p} in (35). Thus, for $k = 1, \cdots, m$, the squared distance $d_E^2(S,S')$ is lower bounded by

$$d_E^2(S,S') \dot{\geq} \mathrm{SNR}^{\delta_{3p,k}(\underline{\alpha}_{3p})} \text{ and } \delta_{3p,k}(\underline{\alpha}_{3p}) = \frac{1}{k}\left[\sum_{i=m-k+1}^{m} (1 - \alpha_{3p,i})\right] - \frac{r}{k} \tag{36}$$

where $\underline{\alpha}_{3p} = [\alpha_{3p,1} \cdots \alpha_{3p,m}]^t$ and $\lambda_{3p,i} = \mathrm{SNR}^{-\alpha_{3p,i}}$.

Remark IV.2. In case the reader ponders over why we have $\frac{2r}{k}$ in (34) (or see below)

$$\delta_{2,k}(\underline{\alpha}_2) := \frac{1}{k}\left[\sum_{i=m-k+1}^{m} (1 - \alpha_{2,i})\right] - \frac{2r}{k},$$

and why in (36) we have $\frac{r}{k}$ for $\delta_{3p,k}(\underline{\alpha}_{3p})$, given both error events $\mathscr{E}_{2,2}$ and $\mathscr{E}_{2,1}$ concern with the case of both users in error, it is simply because of the looseness of mismatch eigenvalue lower bound [27], [24], [23] on $d_E^2(S,S')$ we have used in both cases. The bound is loose in general since almost all of the difference matrices ΔS in $\mathscr{E}_{2,2}$ have determinant $\det(\Delta S \Delta S^\dagger) \gg 1$, and almost all $\det(\Delta X \Delta X^\dagger) \times \det(\tau(\Delta X)\tau(\Delta X)^\dagger) \gg 1$ with $\Delta X = \psi(dx)$ in $\mathscr{E}_{2,1}$. Yet, the algebraic mismatch eigenvalue lower bound captures only the worst case, which actually has probability 0. Furthermore, the difference is also due to the rank of the difference matrix ΔS. To elaborate on this, as the use of mismatch eigenvalue lower bound [27], [24] is closely related to the proof of point-to-point cyclic division algebra based space-time codes being approximately universal [24] for any number of transmit antennas n_t and for any number of receive antennas n_r, below we give a brief insight into that proof, and it will in turn explain why such difference between $\delta_{2,k}(\underline{\alpha}_2)$ and $\delta_{3p,k}(\underline{\alpha}_{3p})$ would occur. Recall in [24], to construct a point-to-point DMT optimal space-time code with multiplexing gain r and with n_t transmit antennas using cyclic division algebra, one of the keys is to set the base-alphabet $\mathscr{B}(\mathrm{SNR})$ as

$$\mathscr{B}(\mathrm{SNR}) = \left\{a + bi : -\mathrm{SNR}^{\frac{r}{2n_t}} \leq a, b \leq \mathrm{SNR}^{\frac{r}{2n_t}},\ a, b \text{ odd}\right\}.$$

Note that it is the same as $\mathscr{A}(\mathrm{SNR})$ of the present construction. Setting $\mathscr{B}(\mathrm{SNR})$ (and the same for $\mathscr{A}(\mathrm{SNR})$) to have size $\mathrm{SNR}^{\frac{r}{n_t}}$ and working on a code matrix of rank n_t give a mismatch eigenvalue lower bound on $d_E^2(S,S')$ with form

$$\delta_k(\underline{\alpha}) \ = \ \frac{1}{k}\left[\sum_{i=m-k+1}^{m}(1-\alpha_i)\right]-\frac{r}{k}$$

as shown in [24]. Error events $\mathscr{E}_1$ and $\mathscr{E}_{2,1}$ are in this category, and hence there is no surprising $\delta_{1,k}(\underline{\alpha}_1)$ and $\delta_{3p,k}(\underline{\alpha}_{3p})$ are of a form similar to $\delta_k(\underline{\alpha})$. Note also that in these cases we have $k=1,2,\cdots,m$, and $m=\min\{n_t,n_r\}$.

The only surprising case is actually $\delta_{2,k}(\underline{\alpha}_2)$ for $\mathscr{E}_{2,2}$, not the others. In $\mathscr{E}_{2,2}$, the difference matrix ΔS has rank $2n_t$. Thus according to the proof in [24], if we want to have a DMT optimal code with rank $2n_t$ and multiplexing gain r, we should set the base-alphabet as

$$\mathscr{B}'(\mathrm{SNR})=\left\{a+bi:-\mathrm{SNR}^{\frac{r}{4n_t}}\leq a,b\leq \mathrm{SNR}^{\frac{r}{4n_t}},\ a,b \text{ odd}\right\}.$$

Note the exponent $\frac{r}{4n_t}$ shown above. But we did not set the base-alphabet as the above in the present construction. Instead, the same base-alphabet $\mathscr{A}(\mathrm{SNR})$ is used in the case of rank being $2n_t$. Note that $\mathscr{A}(\mathrm{SNR})$ can be obtained by $\mathscr{B}'(\mathrm{SNR})$ by replacing the r of $\mathscr{B}'(\mathrm{SNR})$ by $2r$, i.e. $\frac{2r}{4n_t}=\frac{r}{2n_t}$. Thus along the same lines as in [24] we expect the same change from r to $2r$ in $\delta_k(\underline{\alpha})$, i.e.

$$\delta_k'(\underline{\alpha}) \ = \ \frac{1}{k}\left[\sum_{i=m'-k+1}^{m'}(1-\alpha_i)\right]-\frac{2r}{k}$$

and it should be noted that here we have $k=1,2,\cdots,m'$ with $m'=\min\{2n_t,n_r\}$, another difference between $\delta_k(\underline{\alpha})$ and $\delta_k'(\underline{\alpha})$. This is exactly what happened when analyzing the error event $\mathscr{E}_{2,2}$.

Finally we remark that unlike the MAC-DMT proof of Gaussian random codes in [7] where Tse *et al.* used the union bound of pairwise error probabilities for $(2n_t\times T)$ random multiuser codes with SNR^{rT}-fold for the event of one user in error and with SNR^{2rT}-fold for the event of both users in error, here we did not use such argument, i.e. we did not argue using the union bound of pairwise error probabilities. Instead, we argue from the sphere bound of correct decisions, hence the number of nearest neighbors does not come into the scene. The different r's occurred in events $\mathscr{E}_1$, $\mathscr{E}_{2,1}$, and $\mathscr{E}_{2,2}$ are only due to the "mis-setting" of base-alphabet in $\mathscr{E}_{2,2}$. □

It can again be shown similarly that

$$\{\underline{\alpha}_{3p}:\delta_{3p,k}(\underline{\alpha}_{3p})\leq 0,k=1,\cdots,m\}=\left\{\underline{\alpha}_{3p}:\sum_{i=1}^{m}(1-\alpha_{3p,i})^+\ \leq\ r\right\}$$

and that

$$\Pr\{\mathscr{E}_{2,1}\}\ \dot{\leq}\ \Pr\left\{\underline{\alpha}_{3p}:\sum_{i=1}^{m}(1-\alpha_{3p,i})^+\ \leq\ r\right\}=\Pr\left\{\log\det\left(I_{n_r}+\mathrm{SNR}H_{3p}H_{3p}^\dagger\right)\leq r\log\mathrm{SNR}\right\}.$$

To fulfill the second condition required in Theorem IV.4, we need to show

$$\Pr\left\{\log\det\left(I_{n_r}+\mathrm{SNR}H_{3p}H_{3p}^\dagger\right)\leq r\log\mathrm{SNR}\right\}\doteq\mathrm{SNR}^{-d^*_{n_t,n_r}(r)},$$

meaning that at high SNR regime the probability is independent of the choices of p. Be warned that the above is false at low SNR regime, and the probability would depend strongly on p.

To this end, recall that $H_{3p} = H_1 + H_2 P$ and $P = \psi(p)$ and also that for quasi-static Rayleigh fading channel, the entries of H_1 and H_2 are i.i.d. $\mathbf{C}\mathcal{N}(0,1)$ random variables. Let $\underline{h}_{3p,i}^t$ be the ith row of H_{3p}, $i = 1, \cdots, n_r$; then the covariance matrix of $\underline{h}_{3p,i}$ is

$$\Sigma \;=\; E\underline{h}_{3p,i}\underline{h}_{3p,i}^{\dagger} = I_{n_t} + P^t P^*,$$

and $\underline{h}_{3p,i}^t$ and $\underline{h}_{3p,j}^t$ are independent for $i \neq j$. $P^t P^*$ is positive definite since P is invertible in $M_{n_t}(E)$, and hence $P^t P^*$ has the following eigen-decomposition

$$P^t P^* = U^t \Lambda_p U^*$$

for some unitary matrix U; Λ_p is a diagonal matrix whose main diagonal consists of the eigenvalues of $P^t P^*$. Thus, we see

$$\Sigma \;=\; I_{n_t} + P^t P^* = U^t \left(\Lambda_p + I_{n_t}\right) U^* = U^t \Xi U^*.$$

The eigenvalues of Σ are lower bounded by 1, since $\Xi = \Lambda_p + I_{n_t}$. Furthermore, by Karhunen-Loève expansion we see that H_{3p} is statistically equivalent to the matrix

$$G_3 \;=\; G\sqrt{\Xi}U$$

where G is an $(n_r \times n_t)$ random matrix having i.i.d. $\mathbf{C}\mathcal{N}(0,1)$ entries, since both H_{3p} and G_3 have the same joint probability density functions. As a short summary, the above shows

$$\Pr\left\{\log\det\left(I_{n_r} + \mathrm{SNR} H_{3p} H_{3p}^{\dagger}\right) \leq r \log \mathrm{SNR}\right\} = \Pr\left\{\log\det\left(I_{n_r} + \mathrm{SNR} G \Xi G^{\dagger}\right) \leq r \log \mathrm{SNR}\right\}.$$

It should be noted that setting $G_3 \;=\; G\sqrt{\Xi}U$ does not mean the matrix P is known to the receiver at all. We are only saying that the probability $\Pr\left\{\log\det\left(I_{n_r} + \mathrm{SNR} H_{3p} H_{3p}^{\dagger}\right) \leq r \log \mathrm{SNR}\right\}$ can be measured in a different manner.

Now using Minkowski determinant inequality [28] for positive definite matrices which states

$$\left[\det(A+B)\right]^{1/n} \;\geq\; \left[\det(A)\right]^{1/n} + \left[\det(B)\right]^{1/n}, \tag{37}$$

if A and B are $(n \times n)$ positive definite matrices, and for some very small ε, $0 < \varepsilon < 1$ setting

$$A \;=\; (1-\varepsilon) I_{n_r} + \mathrm{SNR} G G^{\dagger} \quad \text{and } B \;=\; \varepsilon I_{n_r} + \mathrm{SNR} G \Lambda_p G^{\dagger},$$

where it should be noted that B is positive definite with probability one (W.P.1), we can show that

$$\begin{aligned}\left[\det\left(I_{n_r} + \mathrm{SNR} G \Xi G^{\dagger}\right)\right]^{1/n_r} \;&=\; \left[\det(A+B)\right]^{1/n_r} \geq \left[\det(A)\right]^{1/n_r} + \left[\det(B)\right]^{1/n_r} \qquad \text{(W.P.1)}\\ &\geq\; \left[\det(A)\right]^{1/n_r} \doteq \left[\det\left(I_{n_r} + \mathrm{SNR} G G^{\dagger}\right)\right]^{1/n_r},\end{aligned}$$

where the last exponential equality follows from $(1-\varepsilon) \doteq \mathrm{SNR}^0$ when $\varepsilon \to 0$. Hence

$$\log\det\left(I_{n_r} + \mathrm{SNR} G \Xi G^{\dagger}\right) \;\dot{\geq}\; \log\det\left(I_{n_r} + \mathrm{SNR} G G^{\dagger}\right)$$

with probability one. Finally we conclude

$$\Pr\left\{\log\det\left(I_{n_r} + \mathrm{SNR} G \Xi G^{\dagger}\right) \leq r \log \mathrm{SNR}\right\}$$

$$\dot{\leq} \Pr\left\{\log\det\left(I_{n_r}+\mathrm{SNR}GG^{\dagger}\right)\leq r\log\mathrm{SNR}\right\} \doteq \ \mathrm{SNR}^{-d^*_{n_t,n_r}(r)}.$$

This completes the proof.

D. An explicit two-user code for $n_t = 2$ and $n_r = 4$ and simulation results

The proposed code is based on the algebra $\mathfrak{D} = (\mathbb{Q}(\xi)/\mathbb{Q}(s),\sigma,\frac{2+i}{2-i})$, where $\xi = \zeta_{16} = e^{\pi i/8},\ \ s = \zeta_8 = \frac{1+i}{\sqrt{2}}$,and $\sigma(\xi) = -\xi,\ \ \tau(\xi) = i\xi,\ \ \tau(s) = -s$. This algebra has been considered before in e.g. [25], [5], [2].

The code for a single user A is (similarly for B)

$$A[a_1,a_2,a_3,a_4,a_5,a_6,a_7,a_8] := (\ A_1,\ \tau(A_1)\ ),$$

where

$$A_1 = \begin{pmatrix} a_1+a_2s+a_3\xi+a_4s\xi & \frac{2+i}{2-i}(a_5+a_6s-a_7\xi-a_8s\xi) \\ a_5+a_6s+a_7\xi+a_8s\xi & a_1+a_2s-a_3\xi-a_4s\xi \end{pmatrix},$$

and

$$\tau(A_1) = \begin{pmatrix} a_1-a_2s+a_3i\xi-a_4si\xi & \frac{2+i}{2-i}(a_5-a_6s-a_7i\xi+a_8si\xi) \\ a_5-a_6s+a_7i\xi-a_8si\xi & a_1-a_2s-a_3i\xi+a_4si\xi \end{pmatrix}.$$

Note that we do not need the minus sign in the right lower corner, as $n_t = 2$.

The code

$$\begin{pmatrix} A_1 & \tau(A_1) \\ B_1 & \tau(B_1) \end{pmatrix}.$$

is DMT optimal. However, as DMT optimality only promises asymptotically good performance, we can add the matrix

$$\Gamma = \begin{pmatrix} \zeta_7 & 0 \\ 0 & \zeta_7 \end{pmatrix}$$

in order to get full rank and hence expectedly good performance also at low and moderate SNRs. That is, at low and moderate SNRs we may use the code

$$\begin{pmatrix} A_1\Gamma & \tau(A_1) \\ B_1 & \tau(B_1)\Gamma \end{pmatrix}.$$

The coefficients a_i, b_i take complex values and the code is orthogonal. By using a maximal order code we could get even better performance, but within this time limit we could not implement the required sphere encoding algorithm (simple PAM or QAM modulation cannot be used with non-orthogonal codes if one wishes to get the advantage provided by the density, see [29] and [4]). In Figure 4 we have plotted the performance of the proposed code (New code), and compare it with the best previously known code by Badr and Belfiore [11].

At the SNR-range of our simulation, it appeared to be irrelevant whether the twist matrix Γ was present or not until quite big SNRs, where totally expectedly the DMT optimal code without Γ clearly starts to win over the full-rank version with Γ. In spite of the fact that adding Γ gives us no gain at low-moderate SNRs, adding (resp. removing) such twist matrix at low-moderate SNRs (resp. high SNRs) to guarantee

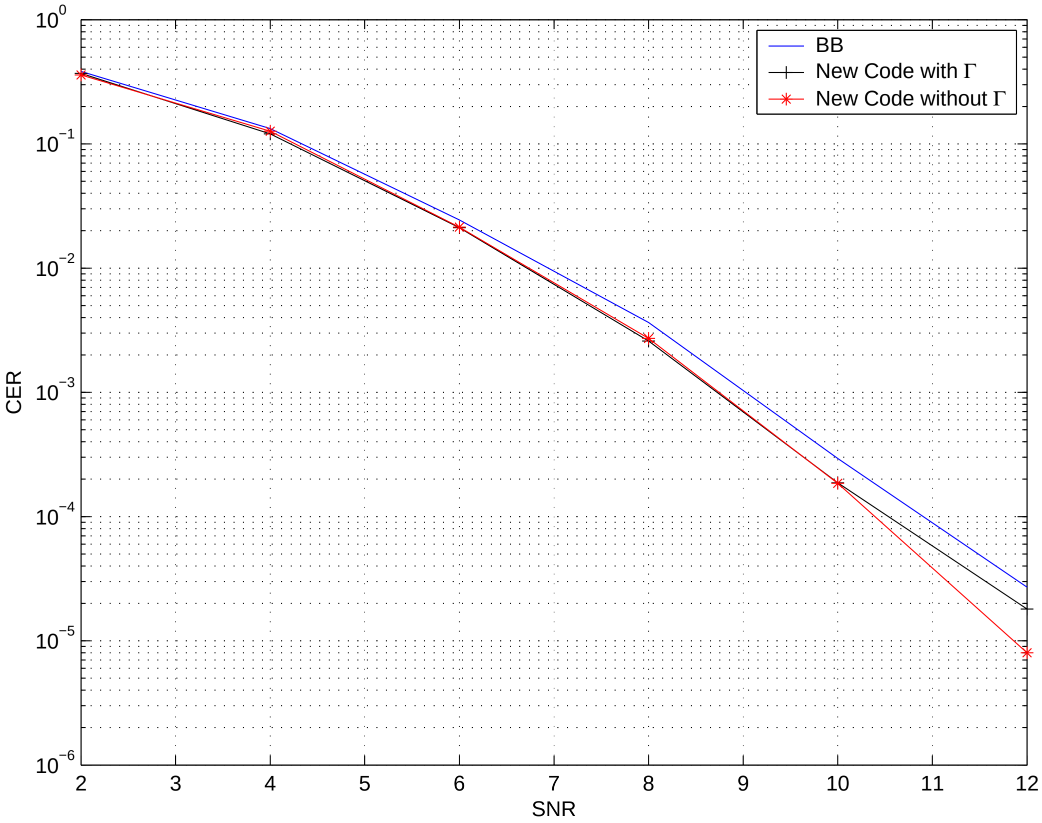

Fig. 4. The performance of the codes on 4-QAM received with 4 antennas.

ful l rank (resp. DMT optimality) may in some other case (different algebra/order/twist matrix) be a useful trick worth checking out.

The proposed code actually has lower density as the code by Badr and Belfiore. Their code has a normalized minimum determinant $1/20$ that gets very close to the upper bound $1/16$ of orthogonal multi-block codes. We have maximal order codes that are denser than that, but as stated above, we did not have a suitable implementation of the sphere decoder in order to simulate them. We also tried the Badr-Belfiore code without the twist matrix Γ they propose, and the performance turned out to be only slightly worse.

Remark IV.3. In case the reader ponders over why we cannot use the above DMT optimal code also for the earlier situation where we have $n_r = 2$, the reason is simply in the decoding: if we wish to use a simple (ML performance preserving) decoding method such as sphere decoding, then receiving the above code calls for at least four receive antennas and hence cannot be efficiently decoded with two receivers only.

Of course we can overcome this by using suboptimal decoders [11, Section V.A], [16] but then we lose the ML performance at least to some extent (cf. Remark I.1).

V. Conclusions

In this paper, we have provided coding schemes based on the multi-block structure. All the codes are sphere decodable, the latter scheme being also MAC-DMT achieving. By computer simulations we have shown that the newly proposed codes outperform the best previously known codes. These satisfactory results were achieved by exploiting cyclic division algebras and their orders to meet the new, relaxed design criteria that were shown to be sufficient for achieving the optimal MAC-DMT.

Appendix I
Nonexistence of Full NVD Multiuser Codes: Proof of Theorem IV.3

Below we will prove the nonexistence of full NVD multiuser codes when each user is equipped with single transmit antenna. Thus, as there are two users in the present case, the overall code matrix is of size (2×2), one row for each user. In the following we show if the (2×2) code matrix has non-zero determinant then it cannot have NVD. We first invoke the following well-known result in lattice theory.

Lemma I.1. *A subgroup in $\mathbf{C}^n$ is a lattice if and only if it is discrete.* □

To prove Theorem IV.3, let us suppose that user one uses a code $\mathscr{C}_1$ that is a full lattice, i.e. it has 4 generators as an abelian group in $\mathbf{C}^2$. The reason for having 4 generators is that the transmission of code takes two channel uses, and in each channel use it is a complex baseband symbol that has two components, the in-phase and quadrature. Let us now suppose that (b_1, b_2) is some non-zero codeword sent by the second user and (a_1, a_2) a nonzero codeword sent by the first user. The two-user matrix is now

$$S = \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}.$$

We have $\det(S) = a_1b_2 - a_2b_1$. Fixing (b_1, b_2) for the second user gives us an idea of a natural homomorphism f from $\mathscr{C}_1$ to $\mathbf{C}$ where $(x_1, x_2) \mapsto x_1b_2 - x_2b_1$. The assumption of S having non-zero determinant for all nonzero $(a_1, a_2) \in \mathscr{C}_1$ suggests that $x_1b_2 - x_2b_1$ is zero if and only if (x_1, x_2) is zero, hence we see that f is a group isomorphism from $\mathscr{C}_1$ to $f(\mathscr{C}_1) \subseteq \mathbf{C}$. Now $f(\mathscr{C}_1)$ is a subgroup in $\mathbf{C}$ and it must have 4 generators as an abelian group because it is isomorphic to $\mathscr{C}_1$. As any lattice in $\mathbf{C}$ can have at maximum 2 generators, we see that $f(\mathscr{C}_1)$ cannot be a lattice. Therefore it must have an accumulation point. Because $f(\mathscr{C}_1)$ is a group we can suppose that it has an accumulation point at 0. This means that there exists an element (a_1, a_2) in $\mathscr{C}_1$ so that we can get $|a_1b_2 - a_2b_1|$ arbitrarily small, yielding a vanishing determinant. Hence this proves there does not exist any multiuser codes that are full NVD.

References

[1] F. Oggier, G. Rekaya, J.-C. Belfiore, and E. Viterbo, “Perfect space time block codes,” *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3885– 3902, Sept. 2006.

[2] P. Elia, B. A. Sethuraman, and P. V. Kumar, “Perfect space-time codes for any number of antennas,” *IEEE Trans. Inf. Theory*, vol. 53, no. 11, pp. 3853–3868, 2007.

[3] H.-F. Lu, “Explicit constructions of multi-block space-time codes that achieve the diversity-multiplexing tradeoff,” in *Proc. 2006 IEEE Int. Symp. Inform. Theory*, Seattle, WA, Jul. 2006, pp. 1149–1153.

[4] K. R. Kumar and G. Caire, "Space-time codes from structured lattices," *IEEE Trans. Inf. Theory*, to appear, 2008. Preprint available at http://www.citebase.org/abstract?id=oai:arXiv.org:0804.1811.

[5] R. Vehkalahti, C. Hollanti, J. Lahtonen, and K. Ranto, "On the densest MIMO lattices from cyclic division algebras," *IEEE Trans. on Inform. Theory*, to appear, 2009. Preprint available at: http://arxiv.org/abs/cs.IT/0703052.

[6] M. E. Gärtner and H. Bölcskei, "Multiuser space-time/frequency code design," in *Proc. 2006 IEEE Int. Symp. Inform. Theory*, Seattle, WA, Jul. 2006, pp. 2819 – 2823.

[7] D. Tse, P. Viswanath, and L. Zheng, "Diversity and multiplexing tradeoff in multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1859–1874, 2004.

[8] L. Zheng and D. Tse, "Diversity and multiplexing: a fundamental tradeoff in multiple antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.

[9] Y. Nam and H. E. Gamal, "On the optimality of lattice coding and decoding in multiple access channels," in *Proc. 2007 IEEE Int. Symp. Inform. Theory*, Nice, France, Jun. 2007.

[10] Y. Hong and E. Viterbo, "Algebraic multi-user space-time block codes for 2x2 MIMO," in *Proc. 2008 IEEE PIMRC*, Cannes, France, Sep. 2008.

[11] M. Badr and J.-C. Belfiore, "Distributed space-time block codes for the MIMO multiple access channel," in *Proc. 2008 IEEE Int. Symp. Inform. Theory*, Toronto, ON, Jul. 2008.

[12] W. Zhang and K. B. Letaief, "A systematic design of multiuser space-frequency codes for MIMO-OFDM systems," in *Proc. 2007 IEEE Int. Conf. Commun.*, Jul. 2007, pp. 1054–1058.

[13] P. Coronel, M. Gärtner, and H. Bölcskei, "Diversity multiplexing tradeoff in selective fading multiple-access MIMO channels," in *Proc. 2008 IEEE Int. Symp. Inform. Theory*, Toronto, ON, Jul. 2008.

[14] R. Vehkalahti, J. Lahtonen, C. Hollanti, and H.-F. F. Lu, "On the code design criteria for MIMO multiple access channels: A non-existence result," submitted to ITW 2009.

[15] F. E. Oggier, J.-C. Belfiore, and E. Viterbo, "Cyclic division algebras: A tool for space-time coding," *Foundations and Trends in Communications and Information Theory*, vol. 4, no. 1, pp. 1–95, 2007.

[16] A. D. Murugan, H. E. Gamal, M. O. Damen, and G. Caire, "A unified framework for tree search decoding: rediscovering the sequential decoder," *IEEE Trans. Inf. Theory*, vol. 52, pp. 933 – 953, March.

[17] I. Reiner, *Maximal Orders*. New York: Academic Press, 1975.

[18] P. Elia and P. V. Kumar, "Approximately-universal space-time codes for the parallel, multi-block and cooperative-dynamic-decode-and-forward channels," Jul. 2007, http://arxiv.org/pdf/0706.3502.

[19] C. Hollanti and J. Lahtonen, "A new tool: Constructing STBCs from maximal orders in central simple algebras," in *Proc. 2006 IEEE Inform. Theory Workshop*, Punta del Este, Uruguay, Mar. 13-17 2006.

[20] J. Liu and A. R. Calderbank, "The icosian code and the e_8 lattice: A new 4×4 space-time code with nonvanishing determinant," in *Proc. 2006 IEEE Int. Symp. Inform. Theory*, Seattle, WA, Jul. 2006, pp. 1006–1010.

[21] C. Hollanti, J. Lahtonen, and H.-F. Lu, "Maximal orders in the design of dense space-time lattice codes," *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4493 – 4510, Oct. 2008.

[22] S. Yang and J.-C. Belfiore, "Optimal space-time codes for the amplify-and-forward cooperative channel," in *Proc. 43nd Allerton Conference on Communication, Control and Computing*, Monticello, Illinois, Sep. 2005.

[23] H.-F. Lu, "Explicit constructions of multi-block space-time codes that achieve the diversity-multiplexing tradeoff," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3790–3796, 2008.

[24] P. Elia, K. R. Kumar, S. A. Pawar, P. V. Kumar, and H.-F. Lu, "Explicit construction of space-time block codes achieving the diversity-multiplexing gain tradeoff," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3869–3884, Sep. 2006.

[25] T. Kiran and B. S. Rajan, "STBC-schemes with non-vanishing determinant for certain number of transmit antennas," *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2984–2992, Aug. 2005.

[26] J.-C. Belfiore, G. Rekaya, and E.Viterbo, "The Golden code: a 2×2 full-rate space-time code with non-vanishing determinants," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1432–1436, Apr. 2005.

[27] C. Köse and R. D. Wesel, "Universal space-time trellis codes," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2717–2727, Oct. 2003.

[28] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge: Cambridge University Press, 1985.

[29] C. Hollanti and K. Ranto, "Maximal orders in space-time coding: Construction and decoding," in *Proc. 2008 Int. Symp. Inf. Theory and its Appl. (ISITA)*, Auckland, New Zealand, Dec. 2008.